Nonlinear electrohydrodynamic phenomena and droplet generation in charged jets of conducting liquid

V. N. Gorshkov and M. G. Chaban

Institute of Physics, Academy of Sciences of Ukraine, 252650 Kiev, Ukraine
(Submitted June 4, 1998; resubmitted October 23, 1998)
Zh. Tekh. Fiz. **69**, 1–9 (November 1999)

Phenomena occurring at the tip of a charged conducting jet are analyzed in detail using numerical methods developed for axially symmetric flows. Universal mechanisms (independent of the method for producing the jet) for droplet formation with different ratios of the Laplace and electrical pressures on the lateral surface are identified. An explanatory analysis is given for all of the nonlinear stages of the classical Rayleigh instability of a charged conducting drop, beginning with the formation of a jet at the surface of the drop and culminating in the generation of a developed jet of secondary droplets. © *1999 American Institute of Physics.* [S1063-7842(99)00111-7]

1. Studies of the physical mechanisms governing the breakup of a cylindrical liquid jet into droplets have a rather long history, beginning with the classical work of Rayleigh.[1] The ideas behind the linear theory of the instability of an infinite jet in the absence of an electric field are perfectly clear. Random small perturbations in its radius of the form $\varepsilon \exp(ikz)$ $(k=2\pi/\lambda)$ disturb the uniformity of the Laplace pressure $p_L=\alpha(1/R_1+1/R_2)$ at the surface (α is the coefficient of surface tension and $R_{1,2}$ are the principal radii of curvature). In the case of long-wavelength disturbances, p_L is determined by the curvature of the surface, $1/R_1$, in the plane perpendicular to the axis of the jet, and the liquid flows from constricted zones toward wider zones, which causes an exponential growth in the initial perturbations. For short-wavelength disturbances ($\lambda<\lambda_k=2\pi r_0$, where r_0 is the initial radius of the jet), the curvature $(1/R_2)$ of a surface passing through the axis ($R_2<0$ in a constricted region) becomes important. The resulting pressure imbalance leads to a reverse flow of liquid, to a reduction in the initial perturbations, and to surface oscillations. With increasing λ ($\lambda>\lambda_k$) the amplitude of the variations in $p_L(z)$ ($\Delta p_L \to 2\alpha\varepsilon/r_0^2$) increases, but at the same time there is an increase in the mass of liquid whose flow over a length equal to λ brings about the aperiodic development of each of the constrictions. The optimum value of λ, corresponding to the maximum growth rate, is given by $\lambda_m \approx 9r_0$.

If a conducting jet is charged, then λ_k and λ_m are smaller, since the negative electrostatic pressure, which is greater in absolute magnitude in regions with a high surface curvature, provides correctives to the pressure drop.[2]

As to the nonlinear stages of the development of the instability, for a long time little was known about the mechanism for droplet and satellite formation, when a jet breaks up at two cross sections to the right and left of the site of the original constriction, rather than at the constriction itself, thereby creating primary droplets with radii on the order of $1.9r_0$, as well as small droplets of size $r_0/3$ (for $\lambda=\lambda_m$).[3] Only a detailed analysis of a numerical simulation revealed the physical nature of this phenomenon.[4]

In technical applications droplets are generated by finite liquid jets. Thus the theoretical analysis of the instability of infinite jets is unsuitable for understanding the droplet formation process and choosing optimal operating conditions for the relevant devices. For example, the dynamics of extended liquid jets bounded on two sides (in the absence of an electric field) is beautiful and unexpected[5] from the standpoint of Rayleigh's classical approach.[1] In a contracting jet, periodic structures with a characteristic size of the order of $2r_0<\lambda_k$ are self-excited! During their capricious evolution, these jets create a series of droplets of different sizes. Breakup of this sort has been examined numerically in a number of papers[6,7] without analyzing the physical mechanisms for the phenomenon. One such analysis has been carried out[8] and it was shown that nonlinear surface waves are excited under the influence of an overpressure $p_L \approx 2\alpha/r_0$ at the ends of the jet (in the main region, $p_L \approx \alpha/r_0$). The peculiar ''resonance'' at a length λ_r of the order of $2r_0$ occurs because local perturbations in the surface pressure are transmitted by the liquid in both directions from the excitation zone with a characteristic damping length $\approx r_0$. The interaction of the counterpropagating waves leads to chaotic fractionation of the corrugated jet into droplets.

This phenomenon necessitates a new way of looking at the development of the Rayleigh instability of a jet flowing out of an orifice (of course, without contradicting the quantitative results of our predecessors). When there are no external interactions, it is customary to assume that the source of the initial perturbations is the thermal noise of the jet radius,[2] of which the harmonic with the highest growth rate γ_m is selected during transport along the flow. We regard this approach as incorrect and offer the following description of the process of droplet formation. The end of the jet, with a Laplace overpressure p_L, excites a corrugation in the surface with a period of $\lambda_r \approx 2r_0$. Nonlinear effects[8] subsequently give rise to the formation of a droplet at the end of

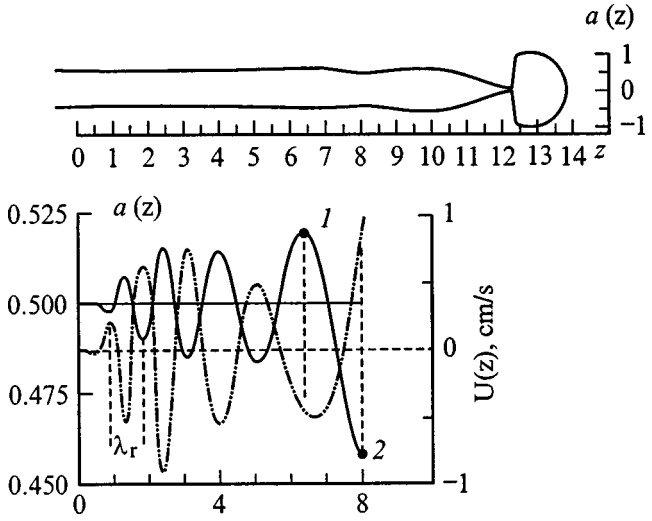

FIG. 1. Generation of droplets from a bounded jet of liquid (in the absence of external perturbations) owing to self-excitation of short-wavelength structures on the surface. Shown here are the full profile $a(z)$ of the perturbed portion of the jet (top) and fragments of the jet radius ($a(z=0)=0.5$ cm) and longitudinal velocity U of the liquid as functions of distance (bottom).

the jet, and a train of waves with the leading (shortest, in accordance with the dispersion relation of Ref. 1) ''resonant'' harmonic $\lambda_r \approx 2r_0$ (Fig. 1) propagates toward the base of the jet. The long-wavelength components that close the wave train impart to the surface a higher level of disturbances, not comparable to the thermal noise. The data of Fig. 1 were obtained from a numerical simulation according to the scheme of Ref. 8 for the dynamics of a cylindrical jet of water, bounded on the right, with a radius $r_0=0.5$ cm (neglecting the force of gravity and with zero initial liquid velocity). The time at which the third droplet breaks away is shown here. At the front of the train, the phases of the oscillations in the radius $a(z,t)$ of the surface and of the average longitudinal velocity $U(z,t)$ are shifted by π, which is typical of a wave process. The change in the phase shift at points *1* and *2* corresponds to growth of the perturbations ($\partial U/\partial z < 0$ at *1* and greater than zero at *2*). The waves are driven on account of the drop in surface energy during the formation of the next droplet. The process has not gone to completion, but the size of the droplet is already close to the observed value $\approx 1.9r_0$. It is difficult to escape the illusion that some perturbations are amplified as they are carried away from the base of the jet. In the steady state the time t_k for the jet to break up (droplets to form) can be estimated by equating the length of the cut-off portion of the jet, λ_m, to the distance the ''resonant'' harmonic propagates over the time t_k, i.e., by setting $\lambda_m = t_k V_{\rm ph}$, where $V_{\rm ph}$ is the corresponding phase velocity of the surface wave. After some calculations using the Rayleigh dispersion relation,[1] we obtain the standard formula $t_k \sim 2/\gamma_m$, in accordance with numerous experimental measurements.

These results will help later in interpreting the physical phenomena observed in our numerical simulations of the breakup of a charged conducting jet. Attempts to model the dynamics of conducting liquids in strong electric fields have been made previously, but the calculations ended at the stage when conical protrusions (Taylor cones) develop.[9,10] Jet regimes were simply unattainable with the numerical schemes employed there.

Based on general physical considerations, we can state that after a jet is formed, the electrohydrodynamic processes near its tip are independent of the manner in which it was produced. In our work we have chosen the instability of highly charged particles as a starting point. At some level of charge, known as the Rayleigh critical charge,[11] a drop becomes unstable and is deformed, ejecting a jet at whose tip droplets are generated. While there has been much theoretical and experimental work on the Rayleigh limit,[12] almost nothing is known about the dynamics of the breakup of a charged drop. Theoretical study of this phenomenon is made difficult by the strong nonlinearity of the process. The energy approach[13] to the problem cannot be used to determine the mechanisms for interesting nonlinear processes. The transient nature of jet breakup in various physical systems (liquid metal ion sources, electrospray devices[14,15]) and the small sizes of these jets make experimental investigation of the essence of many of these processes difficult. Even those phenomena which are detected experimentally are often associated with uncontrollable external interactions. For example, the generation of oblate (along the direction of an external electric field) spheroids, which, in turn, ejected new jets in a transverse direction, has been observed.[16] Far from the tip of the jet, prolate spheroids were observed, along with the generation of secondary droplets in the longitudinal direction. We have shown, in particular, that these are fragments of a single prolonged process, and not the result of aerodynamic effects.[16] The formation of a stepped (with thickening in the transition zone) jet profile (see the plots of $a(z,t)$ in Figs. 4–6) in the experiments of Taylor[17] is a regular stage in the evolution of the balance between the electrical and Laplace pressures along a jet as it grows.

2. For axially symmetric flows of a viscous, incompressible liquid with a free boundary we have used a model that we developed earlier.[8] This model provides a fairly accurate description of the dynamics of even relatively short-wavelength perturbations of the jet, although the longitudinal velocity of the liquid particles, $U(z,r,t)$, was assumed independent of the distance r to the axis. In this case, the radial velocity profile $V(r,z,t)$ is a linear function of r ($V(r,z,t) = V_0(z,t)r/a(z,t)$, where $V_0(z,t)=V(r=a,z,t)=da(z,t)/dt$ is the distribution of the radial component of the velocity of the points on the surface and $a(z,t)$ is the jet profile.), while the corresponding Navier–Stokes equation of motion in a cylindrical coordinate system takes the form

$$(\rho r/a)dV_0/dt = -\partial p_f/\partial r + \mu r \partial^2(V_0/a)/\partial z^2, \qquad (1)$$

where ρ and μ are the density and viscosity of the liquid.

According to Eq. (1), the spatial variation of the pressure $p_f(z,r,t)$ in the liquid is given by

$$p_f(z,r,t) = p_0(z,t) - (\rho r^2/(2a))dV_0/dt + (\mu r^2/2)\partial^2(V_0/a)/\partial z^2, \qquad (2)$$

where $p_0(z,t)$ is the pressure on the jet axis.

Since the pressure p_f at $r=a$ in a given transverse cross section of the jet equals the surface pressure $p_S(z,t)$, we obtain the following equation for the velocity $V_0(z,t)$ from Eq. (2):

$$\rho dV_0(z,t)/dt = 4[p(z,t)-p_S(z,t)]/a(z,t) + \mu a\partial^2(V_0/a)/\partial z^2, \quad (3)$$

where $p(z,t)=(p_0(z,t)+p_S(z,t))/2$.

The force F_z acting on a transverse cross section of the jet is determined by integrating Eq. (2) from 0 to a. Making some transformations using Eq. (3), we obtain $F_z = \pi a^2(z,t)p(z,t)$. Our earlier assumption that $\partial U(z,r,t)/\partial r=0$ signifies an arbitrary division of the liquid into thin disks (truncated cones) whose boundaries have velocities $U(z,t)$. Including all the forces acting on a disk of this sort with thickness Δz leads to an equation for the longitudinal velocity $U(z,t)$,

$$\rho dU(z,t)/dt = -\partial p/\partial z + \mu\partial^2 U(z,t)/\partial z^2 + 2[p_S-p]\partial(\ln a)/\partial z. \quad (4)$$

The right-hand side of Eq. (4) includes terms owing to the gradient in F_z and to the surface force $2\pi a p_S(z,t)\Delta z\partial a/\partial z$ acting on the lateral surface of the disk in the longitudinal direction. Thus Eq. (4) is the standard Navier-Stokes equation of motion including the existence of a free boundary for the given macroscopic particle.

The equation for $p(z,t)$ is easily obtained as follows. We select a jet segment of length Δz and radius $a(z,t)$. When it is deformed during the flow, $a^2(z,t)\Delta z(t)=\text{const}$. Differentiating this equation with respect to time yields

$$2\Delta z(t)da(z,t)/dt + a(z,t)d(\Delta z(t))/dt = 0. \quad (5)$$

Since $d(\Delta z(t))/dt = \Delta z(\partial U/\partial z)$, Eq. (5) yields the equation of continuity for the flow,

$$2V_0(z,t)/a(z,t) + \partial U(z,t)/\partial z = 0. \quad (6)$$

Again, differentiating Eq. (5) with respect to time and using Eq. (6) and the equation $d^2(\Delta z(t))/dt^2 = \Delta z[\partial(dU/dt)/\partial z]$ yield

$$dV_0/dt = 3V_0^2/a + (a/2)[\partial(dU/dt)/\partial z]. \quad (7)$$

After substituting the right-hand sides of Eqs. (3) and (4) in Eq. (7), we obtain a rather cumbersome equation for $p(z,t)$, for which a satisfactory difference approximation that conserves the total volume of liquid is far from obvious. However, such a computational scheme is easily constructed without direct recourse to Eq. (7). (A detailed description of this method is given in Ref. 8.)

The approximation $\partial U(z,r,t)/\partial r=0$ has been used in many papers. It limits the use of the initial system of equations for numerical study of short-wavelength perturbations of a jet. In our work, however, this was the only approximation. In the following we calculate the internal pressure of the liquid and do not profile it in some way, as was done, for example, in Ref. 18: $p_f(z,r,t)=p_S(z,t)$ (after which the equation of motion was solved only for a single component of the velocity and the second component was determined from the equation of continuity). In accordance with Eqs. (2) and (3), we obtain a relationship between the local value of p_f and its average $p(z,t)$ over the transverse cross section of the jet in a natural way,

$$p_f(z,r,t) = p_S(z,t)(2(r/a)^2-1) + 2p(z,t)(1-(r/a)^2) \quad (8)$$

and the equations of motion for each component of the velocity, together with the equation for $p(z,t)$, are solved self-consistently, which ensures conservation of the volumes of each cell (macroscopic particle) of liquid during deformation (during the flow).

This approach extends the range of possibilities for the system of Eqs. (3) and (4) employed here. Thus the dispersion relation for small amplitude surface waves in infinite jets of radius r_0 obtained from the approximate system of hydrodynamic equations has no significant errors compared to an exact solution, even for wavelengths $\approx \pi r_0/2$.[8]

In the cases of low viscosity and slowly varying jet radius, $p(z,t)$ is determined by the steady state diffusion equation

$$\partial^2 p/\partial z^2 - (8/a^2)(p-p_S) + 6\rho(V_0/a)^2 = 0. \quad (9)$$

In our problem, the surface pressure $p_S(z,t)=p_L(z,t)-p_E(z,t)$, where $p_L=\alpha(1/R_1+1/R_2)$, $R_1=a\psi$, $R_2=-\psi^3(\partial^2 a/\partial z^2)^{-1}$, $\psi=[1+(\partial a/\partial z)^2]^{1/2}$, $p_E(z,t)=E^2(z,t)/8\pi$, and $E(z,t)$ is the electric field strength at the surface.

In the case of an ideally conducting liquid, the charge density distribution is found from the condition that the electrical potential be constant at all points on the surface.[19] The liquid was entirely broken up into a set of truncated cones (with a small ratio of the cone height to the radius of its average cross section) with their planes perpendicular to the axis of the flow. The charge densities σ_i on the lateral surface of each of these cones of area s_i are determined from the system of equations $\Sigma b_{ki}\sigma_i + \varphi_k = V(t)$ and $\Sigma S_i\sigma_i = Q$ if the flow dynamics is considered to have a specified charge Q ($V(t)$ is the surface potential and φ_k is the potential created by the external field on the surface of cone number k). In calculating the coefficients b_{ki}, the charge on the ith cone is usually represented by a system of point charges located on the middle line of the side surface of the cone, which simplifies the averaging procedure. In our case, where we are solving the stability problem and highly accurate field calculations are necessary, this approximation is unsuitable (especially for calculating the diagonal elements b_{kk}). In the exact expression for b_{ki} an analytic integral is taken with respect to z and a partially numerical and partially analytic integral with respect to the azimuth. This computational method ensured a relative error in the determination of the charge density of less than 0.01% compared to the known analytic solutions. We supplement Eqs. (1)–(3) with the initial and boundary conditions. At $t=0$ a droplet of radius r_0 with its center at the coordinate origin $r=0$, $z=0$ is deformed into an ellipsoid of revolution that is prolate along the Z axis with semiaxes $r_0(1+\delta)$ and $r_0(1+\delta)^{-1/2}$, where $\delta=0.05$. The initial velocities $U(z,t=0)=V_0(z,t=0)=0$ and the total electric charge on the surface equals Q. The boundary conditions are

$$\partial V_0/\partial z|_{z=0} = \partial p/\partial z|_{z=0} = U(z=0,t) = 0,$$

$$a(z_c,t) = 0, \quad p(z_c,t) = p_S(z_c,t), \tag{10}$$

where $z_c(t)$ is the coordinate of the end of the jet.

At each time step the electrostatic problem was solved to determine $p_E(z)$ and $p_L(z)$ was calculated for the given $a(z)$. Then the function $p(z)$,[8] was determined using Eqs. (3), (4), (7), and (10). After that, the new velocities and coordinates $a(z,t+\Delta t)$ were determined using the same Eqs. (3) and (4). As the droplet deforms, the coordinate grid is periodically realigned to ensure the required accuracy of the model for the jet formation and dynamics.

Note that the numerical model for the hydrodynamic part of the problem comes well recommended for studies of the dynamics of self-excited short-wavelength periodic structures in bounded jets.[8] When we included the electrostatic part, as a test we calculated the small oscillations of a droplet with a subcritical charge. The resulting agreement with the analytic solution[11] for the period of the oscillations offers the hope that the calculations will be reliable in more complicated cases.

For small deformations of the droplet, which are described using the associated Legendre polynomials, the first of the modes becomes unstable for $Q>Q_R=(16\pi\alpha r_0^3)^{1/2}$, the Rayleigh limit in the absence of an external electric field for $n=2$ in the dispersion relation[11]

$$\omega^2 = n(n-1)[(n+2)\alpha - Q^2/(4\pi r_0^3)]/(\rho r_0^3). \tag{11}$$

In our calculations we shall change the supercriticality parameter δ_e, defined by $Q=(1+\delta_e)Q_R$. For a given δ_e, the initial system of Eqs. (3), (4), and (6) is reduced to dimensionless form by introducing the new variables $t'=t(\alpha/(\rho r_0^3))^{1/2}$, $z'=z/r_0$, $a'=a/r_0$, and $\mu'=\mu/(\rho\alpha r_0)^{1/2}$. Thus, all of the many solutions of the problem are determined by two parameters: δ_e and μ'. For simplicity, we have assumed that $\rho=1$ g/cm^3, $r_0=0.5$ cm, and $\mu=1$ gs^{-1}cm^{-1}, and varied δ_e and α in the calculations.

3. Figure 2 shows the results of a numerical simulation with a fairly substantial excess charge on the droplet $Q_R(\delta_e=0.3,0.5)$. In this case the dimensions of the jet of secondary droplets are comparable to those of the initial drop, so it is easier to analyze the electromagnetic phenomena. In Fig. 2 and below, the pressure is given in g/(s^2 cm) and the velocity, in cm/s, while the profiles of the surfaces are given without distorting the ratio of the longitudinal and transverse dimensions. Two features of the process by which a drop breaks up can be seen easily in Fig. 2: (1) the region in which the jet develops (the distance from the minimum radial velocity to the end of the jet; curves *2–4* of Fig. 2a) narrows abruptly in time, and (2) the end of the jet is far from spherical in shape and its cutoff is accompanied by the formation of an oblate secondary droplet.

Equation (9) shows that the depth to which the negative overpressure penetrates into the depth of the liquid at the end of the jet (Fig. 2a, inset), where $p_S(z)>p_S(z_c)$ on the surface, is proportional to some effective value of a that depends on the shape of the surface near $z\sim z_c$. A narrowing of the jet with acceleration takes place only in the region

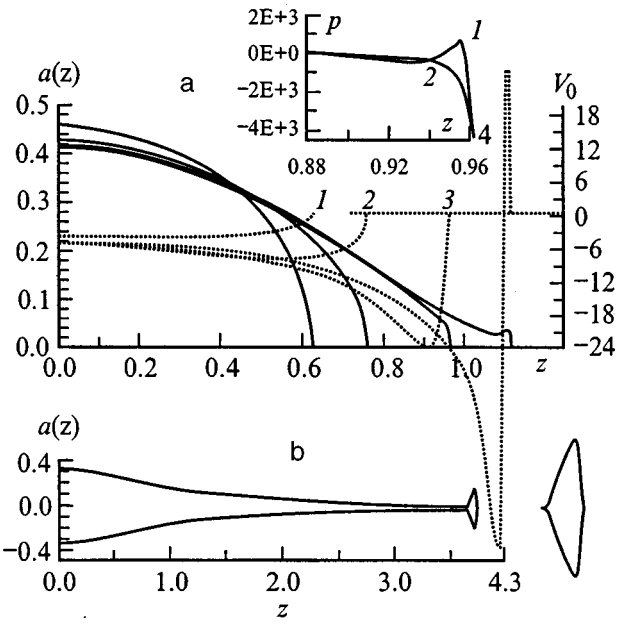

FIG. 2. a: Profiles of the droplet surface and the radial velocity (curves *1–4*) for $\delta_e=0.3$; $\alpha=300$ g/s^2; $t=0.0257$, 0.0319, 0.0354, 0.0363 s. Inset: *1* — $p_S(z)$, *2* — $p(z)$ at $t=0.0354$ s. b: $\delta_e=0.5$, $\alpha=75$ g/s^2. The initial droplet shape at the time the droplet/leader breaks away, $t=0.0799$ s. (A magnified view of the droplet is on the right.) Droplet radius for a spherical shape, $r_k=0.1$ cm, charge on the droplet $q_k=0.085Q$.

with $p(z)-p_S(z)<0$. Elongation (tapering) of the tip reduces the penetration depth of the increasing (in absolute value) $p_S(z_c)$. In addition, the negative pressure is squeezed out from the depth of the jet ($z<z_c$) toward its end by the pressure ''sources'' $\propto V_0^2/a^2$ in Eq. (9). As a result, the cutoff zone for the secondary droplet (the minimum in V_0) swiftly approaches the end of the jet, while the cutoff velocity increases sharply (Fig. 2a).

The formation of an oblate droplet seems strange only at first glance. But for a spherical or ellipsoidal (prolate) end of the jet, an even thinner jet would extend from its tip, and so on, until the size of the secondary droplets reached zero. On the other hand, for $Q>Q_R$ configurations in the form of an oblate ellipsoid of revolution are stable against small axially symmetric deformations.[20,21] Since any external fields will deform a spherical droplet into an prolate ellipsoid, under real conditions with $Q>Q_R$, according to Basaran's hypothesis,[21] oblate spheroids cannot be observed. However, as we shall see, during the breakup of a jet such configurations can develop naturally if the charge on the secondary droplet exceeds a critical value. Let us follow the dynamics of a jet after the lead droplet breaks away (Fig. 2b). The screening of the end of the jet by the charge of the first droplet causes the Laplace pressure p_L, which tends to shorten the jet, to dominate in this zone. Under these conditions, the already familiar mechanism exciting surface corrugation operates.[8] The region where the nucleus of the second droplet adjoins the main part of the jet ($R_2<0$) becomes a zone with a reduced pressure p_S (inset to Fig. 3). The flow of liquid into this zone from the jet leads to the development of a constriction[1)] (inset to Fig. 3) and to the simultaneous development, at its left, of a new region with an elevated pres-

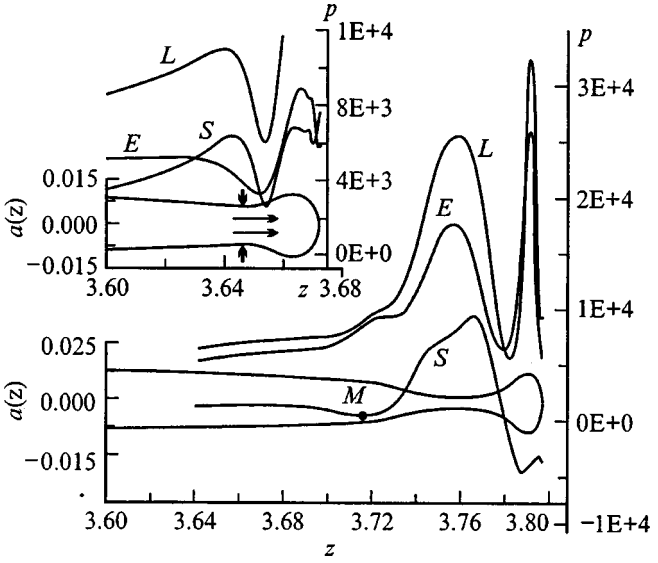

FIG. 3. Self-excitation of corrugation on the jet surface after breakaway of the droplet/leader (the continuation of the variant of Fig. 2b). Profiles of the surfaces and pressures $p_L(z)$, $p_E(z)$, and $p_S(z)$ (the corresponding subscripts are indicated beside the curves) for $t=0.0811$ and 0.0803 s (inset).

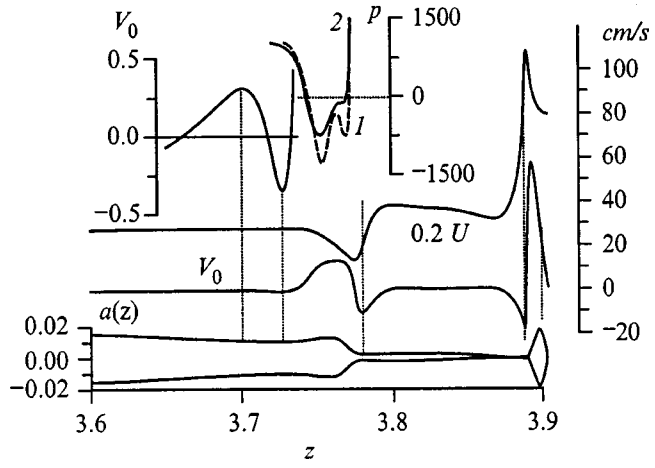

FIG. 4. Spatial variations in the characteristics of the jet (continuation of Fig. 3) at the time a second droplet breaks away: $t=0.0815$ s, $r_k=0.012$ cm, ratio of the charge q_k to the corresponding critical value q_R equals 1.277. The inset above shows fragments of $V_0(z)$, $p_S(z)$ (*1*) and $p(z)$ (*2*).

sure p_S (the next element of the corrugation or the next period of the surface wave) owing to the development of a curvature $1/R_2$.[8] If the charge on the surface were neglected, the process would proceed as shown above (Fig. 1). In this case (the lower part of Fig. 3), however, the screening of the tip of the jet (emitter) decreases as the lead droplet moves away, and this causes an influx of charge. The pressure $p_E(z_c)$ that draws out the jet increases. A second droplet rushes in behind the first and the electric field in the zone where the crest of the surface wave develops (in the future, the third droplet) increases so much as a result of the redistribution of the charge, that a minimum appears in the profile of the total pressure p_S (point M in Fig. 3). The influx of liquid into this zone rises and the amplitude of the crest increases. This leads to further charge accumulation and a negative pressure p_S (inset to Fig. 4). Thus the initial corrugation generated in the course of nonlinear wave processes transforms into an aperiodic instability regime.

As the jet evolves, the characteristic stepped surface profile, which is observed in experiments,[17] develops (Fig. 4). The necks joining the second and third droplets adjoin regions of reduced pressure p_S to their left and right. This situation ends, as is known,[8] in the breakup of the jet at two points (see the $V_0(z)$ profile in Fig. 4; $z_{1,2}\approx 3.78, 3.88$ cm) and the formation of a thin, elongated satellite. Naturally, it will also break up, in accordance with the above scenarios, and generate a multitude of microscopic droplets.

We did no further calculations of the process, but the formation of the next droplet is already noticeable in the radial velocity profile (inset to Fig. 4). The secondary droplets will approach a spherical shape as the residual charge on the initial droplet (the electric field at the tip of the emitter) decreases with time.

Droplet formation is, therefore, entirely determined by nonlinear electrohydrodynamic phenomena at the end of the jet. Previously formed, charged droplets affect the nucleation of new droplets through their electric fields. Thus we should expect some scatter in the sizes of the main droplets and the satellites. (Microscopic droplet chaos of this sort has been observed in liquid metal ion sources.[22])

In the version we examine here (Figs. 2–4), the relatively large δ_e and viscosity of the liquid (parameter μ') caused initial ejection of a jet whose size was such that, in its main part (outside the droplet formation region), the Laplace pressure exceeded the electrical pressure (Fig. 3). For small supercriticalities δ_e and large α (smaller μ'), a thinner jet develops, so that the ratio of p_L and p_E at its lateral surface changes: $p_L<p_E$. In this case, after part of the charge is ejected, we return to the previous (softer, so to say) droplet formation regime. But the initial stages of the breakup of the jet have a somewhat different mechanism for self-excitation of the corrugation in the surface (Fig. 5). Before discussing the physical bases of this phenomenon, let us make a simple mathematical analysis of the topological properties of the function $p_S(z)$ in that part of the liquid which can be called a jet. In the variant of Figs. 2b and 3, a function that rises on the left ($dp_L/dz>0$, since $dR_1/dz<0$) ends on the right with a region where it falls ($dp_S(z)/dz<0$), the electrical charge accumulates, and the electrical pressure dominates. In the intermediate region, $p_S(z)$ should have N maxima and $N-1$ minima. For $N=1$ (the ''fundamental'' mode), the highly charged tip of the jet with $p_S<0$ is cut off (for the profile of Fig. 2b, to the left of the droplet/leader $p_S>0$). The modes $N=2$ and above correspond to cutoff of the droplet with subsequent excitation of a surface wave (to the left of the cutoff zone) by the main pressure jump (inset to Fig. 3, $N=2$).

In the variant shown in Fig. 5 a fairly protracted zone develops in which $p_S(z)<0$. From the standpoint of topology, there is no prohibition on the formation of a pressure maximum $p_S(z)$ to the left of this zone. This sort of pressure jump actually does develop in the later stages of evolution of the jet and ''threatens'' to cut it off at the base as a whole. To generate droplets from the jet, itself, $\max p_S(z)$ must occur in the region where $dp_S(z)/dz<0$ at the left and right

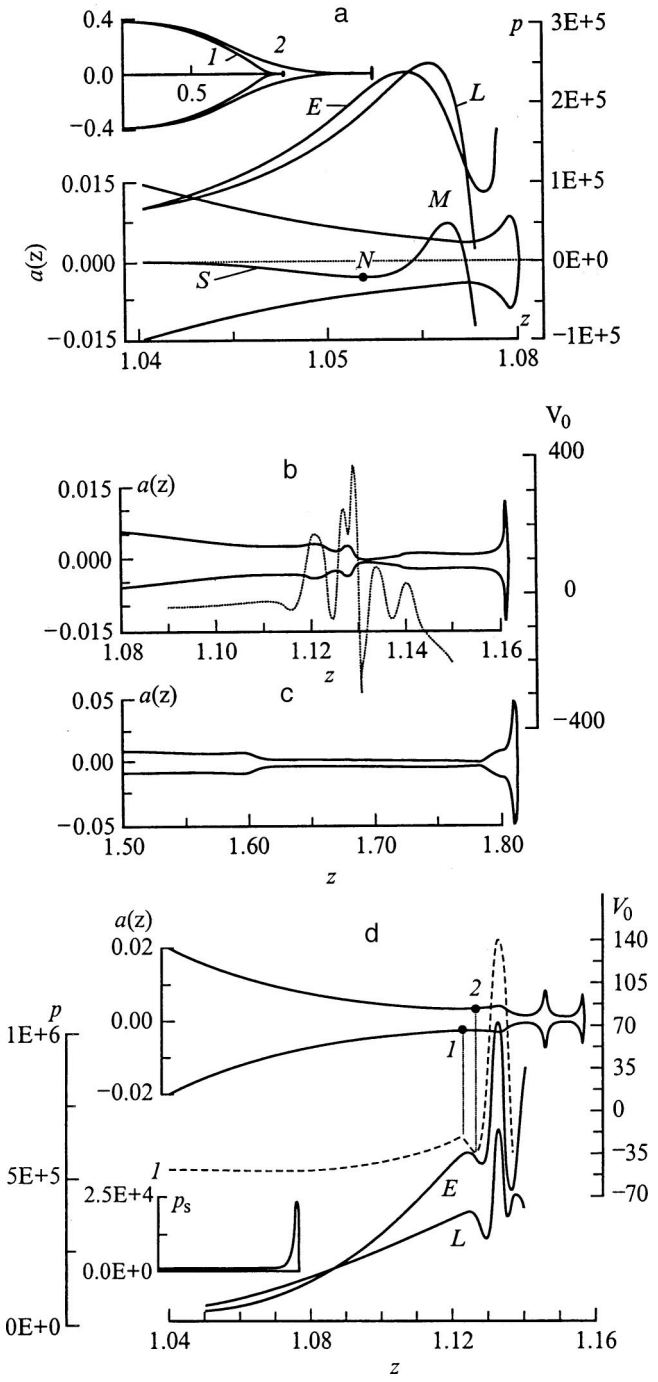

FIG. 5. Illustrating the mechanism for excitation of corrugation in the jet surface for small δ_e. a: — Profiles of the surface and pressures for $\delta_e = 0.05$, $\alpha = 1200$ g/s^2 at $t = 0.06177$ s (the inset shows the overall form of the $a(z)$ curves shown in Fig. 5b and 5c); b: continuation of Fig. 5a: $a(z)$, $V_0(z)$ (dotted curve) for $t = 0.06181$ s; c: $a(z, t = 0.1508\,\text{s})$, $\delta_e = 0.08$, $\alpha = 75$ g/s^2; d: continuation of the variant of Fig. 5b: profiles of the jet surface, $p_L(z)$, $p_E(z)$, and $V_0(z)$ (dotted curve) $\Delta t = 1.7 \times 10^{-5}$ s after breakaway of the head of the jet at the point $z = 1.13$ cm (Fig. 5b). The inset shows $p_S(z)$ within the interval [0;1.09].

ends. Under these conditions, however, extrema can develop only in pairs (max and min, Fig. 5a). Thus a constriction which cuts a droplet off from the tip of the jet (point M) is inevitably accompanied by a broadening of the flow channel (point N)– a prototype of the next (second) droplet. Physically, the extrema in the pressure $p_S(z)$ develop in the following way. If the values of p_L depend only on the local curvature of the $a(z)$ profile, then, besides this factor, the magnitude of the electric field is determined by the distribution of the charge over the entire surface of the deformed droplet. In Fig. 5a, the reduction in p_E on approaching the end of the jet (where a negative curvature $1/R_2$ appears) begins earlier than the reduction in the pressure p_L because of the screening effect of the charge at the tip, which is increasing with time. The shift in the peaks of p_E and p_L creates the two extrema in $p_S(z)$.

The lowered pressure $p_S(z)$ in the region of point N (like the Laplace overpressure in the variant examined before; Fig. 3) initiates self-excitation of corrugation in the surface in the direction of the base of the jet. The formation of a second drop causes electrical charge to accumulate on its surface and a zone with negative curvature $1/R_2$ to form to the left of it. Conditions for formation of a new pair of extrema in $p_S(z)$ are created, etc.

In a later stage of breakup, a distinctive surface profile is observed (Fig. 5b). In this experimental situation it is difficult not to attribute this profile to some uncontrolled perturbations, but, as we have seen, it is the natural evolution of a charged jet. If we increase the viscosity parameter μ', then with the same mechanism for the initial corrugation, the jet profile will be smoother (Fig. 5c).

Figure 5d shows the characteristics of the jet $\Delta t = 1.7 \times 10^{-5}$ s after it has broken off at the point $z = 1.13$ cm of Fig. 5b. Points 1 and 2 correspond to the fourth min and max pair in $p_S(z)$. The formation of the following constrictions and crests can be seen in the $V_0(z)$ profile. The ''principal'' maximum of the pressure $p_S(z)$ has formed near $z \approx 1.07$ cm (inset). This is a transition region between droplet and jet, where the liquid accelerates in the radial (toward the axis) and longitudinal directions before entering the jet. Subsequently, the loss of charge during generation of secondary droplets causes a reduction in the electrical pressure on the surface of the jet and a contraction of the zone where $p_S(z) < 0$. Droplet formation enters the regime shown in Figs. 2b, 3, and 4. In the concluding stage, the jet may be cut off as a whole in the transition region (a similar effect develops in Fig. 4, $z = 3.78$ cm).

If a constant potential is maintained on a jet, then after a large number of droplets have been generated in the interelectrode gap, the electric field on the jet will decrease owing to its being screened by secondary droplets. Periodic breakoff of the jet from the Taylor cone in liquid metal ion sources produces low frequency oscillations in the ion current.[22] As for the ion generation process in these devices, based on the results obtained here (Figs. 4 and 5), it is difficult to conceive that it can take place from a ''hemispherical tip,'' as assumed in a number of papers on the theory of these sources.[14,22]

The possible effect of an electric field on the coefficient of surface tension has been neglected in our calculations. In the most important region for droplet formation (Fig. 5d), $p_E > p_L$. Corrections to the Laplace pressure do not change the physical essence of these phenomena.

Any local pressure perturbations in the region where surface waves are excited (Figs. 3 and 5) will be damped over a

distance Δz from the localization region in proportion to $\exp(-3|\Delta z|/r_c)$ (r_c is the jet radius) according to Eq. (9); that is, they hardly exist for $\Delta z \approx \pm r_c$. During self-organization of the structure of the liquid flow and the surface pressure profile, elements with a characteristic size $\approx 2r_c$ appear, which have been broken off from the tip of the emitter by the strong electric field. Thus in the variants shown in Figs. 4 and 5d, the droplet radius r_k in the initial generation period is of the order of the jet radius ($r_k/r_0 \approx 2.4\times10^{-2}$ and 6×10^{-3}; $q_k/Q \approx 1/300$, $1/700$; $q_k/q_R \approx 1.3$ and 3, where q_R is the critical charge for a spherical droplet shape). As the electric field at the jet surface decreases, the scenario for droplet formation approaches the regime shown in Fig. 1 ($r_k \approx 1.8r_c$).

Let us note yet another feature of droplet formation. As we saw above, the secondary droplets are in the form of oblate spheroids with electrical charges above the critical level. These configurations are unstable with respect to small deformations into a triaxial ellipsoid.[23] Thus we present the scenario for the evolution of the secondary droplets as follows. Microscopic jets develop in the equatorial part of the oblate spheroids and new droplets are generated. This process is possible for the lead droplets in Figs. 4 and 5 until breakoff from the initial droplet. If, on the other hand, the excess charge is not too high (the instability growth rate is low), then breakup of the secondary droplets (with ejection of jets in a direction perpendicular to the direction of motion) will be observed near the tip of the initial jet. As the microscopic droplets are generated, the droplet charge falls below critical. The radial compression of the droplet owing to the Laplace pressure transforms it (because of inertia) into a prolate ellipsoid. Under these conditions, a third stage of jet formation is possible (in a zone further from the emitter), again in the longitudinal direction, even if $q_k < q_R$.[24] This type of dynamics for multistep droplet generation has been reported before, but the observed breakup of oblate and prolate ellipsoids was attributed to random aerodynamic effects.[16,25]

As an illustration of the third stage of droplet formation let us consider the breakup of an uncharged spherical droplet in an external electric field that is below critical $E = 0.9E_T$. (The critical field is[24] $E_T = 1.625(\alpha/r_0)^{1/2}$.) As it stretches out along the field, the droplet passes the equilibrium position as a result of inertia and is sufficiently deformed for instabilities to develop at its tips (Figs. 6a and b).

In Figs. 4, 5b, and 6b, characteristic $p_S(z)$ and $V_0(z)$ profiles with three extrema develop in the regions where the crest of the surface wave has developed ($p_S(z) < 0$, $V_0 > 0$). Figure 6c shows a typical structure of the spatial distributions of the electrical and Laplace pressures in this kind of zone. As the crest develops, $p_E(z)$ and $p_L(z)$ increase at its peak and decrease to the right and left because of the increasing negative curvature of the surface, $1/R_2$. As a result, the initial minimum in the total surface pressure p_S, which stimulates the development of the crest, splits into two minima. (See Fig. 6c; see also around point *M* in Fig. 3 and the inset with a fragment of $p_S(z)$ in Fig. 4.) Note that the increase in the Laplace pressure at the peak of the crest precedes the increase in the electrical pressure and limits the development of surface instabilities in the nonlinear stages. As these calculations show, even before they break away from the jet, the oblate spheroids formed at the tip of the emitter undergo small, axially symmetric (according to our model) oscillations about some equilibrium position.

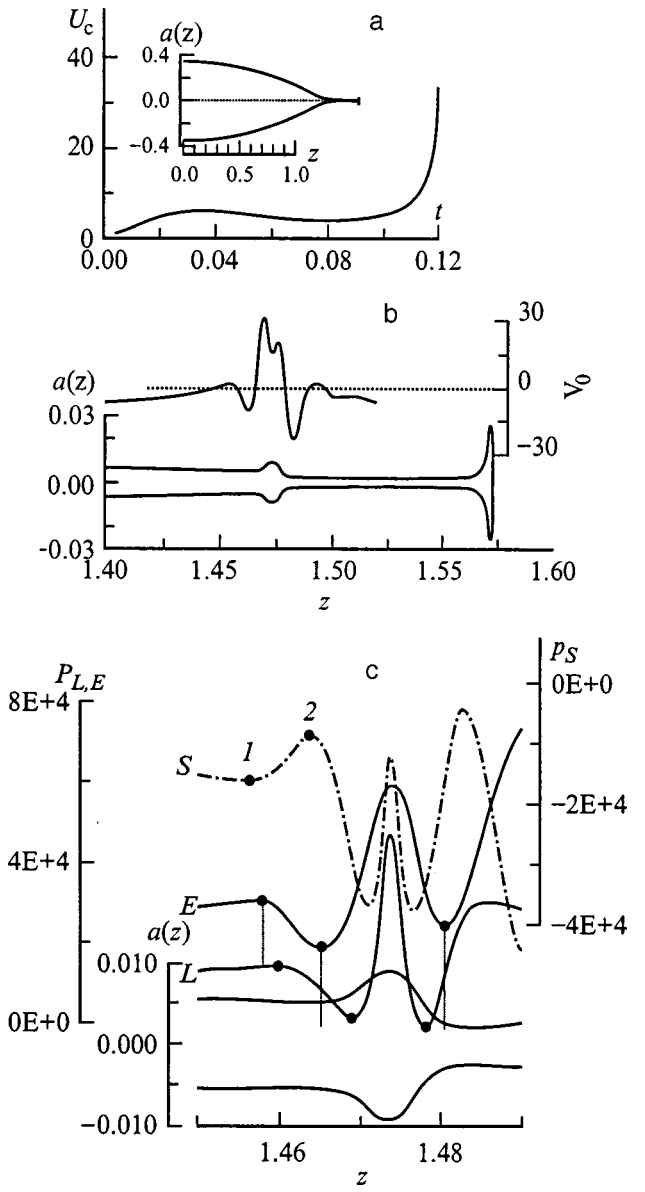

FIG. 6. Breakup of an uncharged droplet in an external electric field. $E = 0.9E_T$, $\alpha = 75$ g/s^2. a: Time dependence of the velocity of the jet tip $U_c = U(z_c, t)$; the inset shows $a(z, t = 0.1228\ \text{s})$; b: fragments of the surface profile and radial velocity near the tip at $t = 0.1228$ s; c: fragments of $p_L(z)$, $p_E(z)$, $p_S(z)$ and $a(z)$ in the region where the jet widens substantially ($t = 0.1228$ s). The extrema of $p_S(z)$, *1* and *2*, are the zones where constrictions on the jet surface and the subsequent crest of the wave develop (see the $V_0(z)$ profile in Fig. 6b).

Droplet generation from a jet of highly conducting, charged liquid is, therefore, determined by a complex of nonlinear electrohydrodynamic processes: excitation of nonlinear surface waves in relatively low electric fields at the tip of the emitter and of short-wavelength aperiodic instabilities with the formation of pairs of extrema in the total surface pressure in strong fields, formation of secondary droplets in the form of oblate ellipsoids, formation of extended satellites

(which break up into a series of microscopic droplets following the same scenario as for the main jet), and stepwise breakup of the secondary droplets through the generation of microscopic jets perpendicular and parallel to the direction of motion (the direction of the external electric field).

[1)]The development of a constriction under these conditions is an essentially nonlinear effect. According to the linear theory, for short-wavelength perturbations the surface pressure increases monotonically with z on going from the zone of the minimum jet radius as the jet expands, and a constriction cannot develop. However, because of nonlinear effects in the formation of the $p_L(z)$ profile, in our case the minimum p_S lies to the right of the constriction.

[1] Lord Rayleigh, Proc. London Math. Soc. **10**, 4 (1878).
[2] S. S. Nazin, A. N. Izotov, and V. B. Shikin, Dokl. Akad. Nauk SSSR **285**, 121 (1985) [Sov. Phys. Dokl. **30**, 606 (1985)].
[3] A. F. Ginevskiĭ and A. S. Dmitriev, in *Physicotechnical Problems in the Study of Monodisperse Systems*, Collected Scientific Papers, No.149 [in Russian], MÉI, Moscow (1987), pp. 5–24.
[4] V. V. Vladimirov and V. N. Gorshkov, Zh. Tekh. Fiz. **60**(11), 197 (1990) [Sov. Phys. Tech. Phys. **35**, 1349 (1990)].
[5] Ya. E. Geguzin, *Droplets* [in Russian], Nauka, Moscow (1977), 176 pp.
[6] N. Tjahjadi, H. A. Stone, and J. M. Ottino, J. Fluid Mech. **243**, 297 (1992).
[7] N. N. Mansour and T. S. Lungren, Phys. Fluids A 1990. **2**, 1141 (1990).
[8] V. N. Gorshkov and D. V. Mozyrskiĭ, Zh. Tekh. Fiz. **66**(10), 15 (1996) [Tech. Phys. **41**, 980 (1996)].
[9] C. Zheng and T. Linsu, J. Vac. Sci. Technol. B **6**, 2104 (1988).
[10] J. D. Sherwood, J. Fluid Mech. **188**, 133 (1988).
[11] Lord Rayleigh, Philos. Mag. **14**, 184 (1882).
[12] E. J. Davis and M. A. Bridges, J. Aerosol Sci. **25**, 1179 (1994).
[13] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61**(3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].
[14] T. Hiramoto, K. Hirakawa, and Y. Ikoma, J. Vac. Sci. Technol. **6**, 1014 (1988).
[15] M. Cloupeau and B. Prunet-Foch, J. Aerosol Sci. **25**, 1021 (1994).
[16] A. Gomez and K. Tang, Phys. Fluids **6**, 404 (1991).
[17] G. Taylor, Proc. R. Soc. London, Ser. A **313**, 453 (1969).
[18] X. D. Sci, Michael P. Brenner, and Sidney R. Nagel, Science **265**, 219 (1994).
[19] A. Renau, F. H. Read, and J. K. Brunt, J. Phys. E **15**, 453 (1982).
[20] J. A. Tsamopoulos, T. R. Akylas, and R. A. Brown, Proc. R. Soc. London, Ser. A **401**, 67 (1985).
[21] O. A. Basaran and L. E. Scriven, Phys. Fluids A **1**, 795 (1989).
[22] V. E. Badan, V. V. Vladimirov, V. N. Gorshkov, and I. A. Soloshenko, Zh. Tekh. Fiz. **63**(6), 47 (1993) [Tech. Phys. **38**, 457 (1993)].
[23] R. Natarajan and R. A. Brown, Proc. R. Soc. London, Ser. A **410**, 209 (1987).
[24] G. Taylor, Proc. R. Soc. London, Ser. A **280**, 383 (1964).
[25] O. A. Basaran and L. E. Scriven, Phys. Fluids A **1**, 799 (1989).

Translated by D. H. McNeill

Instability of a bubble in a liquid dielectric in an external electrostatic field

A. I. Grigor'ev, S. O. Shiryaeva, and A. N. Zharov

P. G. Demidov Yaroslavl State University, 150000 Yaroslavl, Russia
(Submitted June 26, 1998)
Zh. Tekh. Fiz. **69**, 10–13 (November 1999)

The sizes, charges, and number of daughter bubbles emitted during the development of instability with respect to the polarization charge in a uniform electrostatic field of a gas bubble in a liquid dielectric are found on the basis of the Onsager principle of minimum energy dissipation for nonequilibrium processes. © *1999 American Institute of Physics.* [S1063-7842(99)00211-1]

INTRODUCTION

The instability of bubbles in dielectric liquids in strong electric fields is of interest in various areas of applied physics, in particular, for the theory of breakdown in liquid dielectrics. The experimental study of electrical breakdown in liquid dielectrics has been quite complete. Observations show that breakdown is preceded by the formation, near the cathode, of a vapor-gas microbubble which subsequently grows. In a uniform external electrostatic field the bubble loses its spherical shape and is drawn out along the field to form a figure close to a prolate spheroid.[1,2] In a sufficiently strong electric field, the spheroidal bubble becomes unstable: emitting protuberances develop at its ends and small daughter bubbles begin to be ejected from them and carry away the excess charge.[3] Photographs show that at the time the daughter bubbles are ejected, the shape of the bubble is very close to that of a prolate spheroid and that the parent bubble emits several tens of daughter bubbles, which form two clusters, one at each of the emitting protuberances.[3] The daughter bubbles have linear dimensions two orders of magnitude smaller than the bubble which became unstable.

The breakup of bubbles in strong electric fields has been studied very little theoretically. In this connection, the theoretical study of the dispersion of a bubble in a liquid dielectric in an external electric field is of considerable current interest. The present paper is devoted to this problem.

1. We consider an initially spherical bubble of radius R_0 formed during local electrical breakdown in a liquid dielectric and filled with a discharge plasma. Let the liquid have a high heat capacity (so we can neglect the change in the temperature of the system during the instability) and let there be a uniform electrostatic field of strength $\mathbf{E}/\varepsilon$ in it, where ε is the dielectric permittivity of the liquid. Under the influence of the electric field $\mathbf{E}$ the bubble elongates along $\mathbf{E}$ to produce a shape similar to a spheroid of rotation, thereby increasing its volume so that the radius of an equally large sphere increases to R. The charged particles that fill the bubble at the initial time, with their different signs, partially recombine and partially settle on the bubble walls, thereby causing its polarization in the field $\mathbf{E}$. Under the influence of the component of the electrostatic field tangential to the bubble surface, the charges that settle on the walls will be redistributed over the spheroidally deformed surface and cause the charge concentration at the tips to increase.

The characteristic time for balancing of the electrical potential on the bubble surface is given by $\tau_q=\varepsilon R/(\chi E)$, while the characteristic instability development time is $\tau_u=[\rho R^3/\sigma(1-w/(16\pi))]^{1/2}$ (Refs. 4 and 5). In these formulas, $w=E^2R/(\varepsilon\sigma)$ is the Taylor parameter, which characterizes the stability of the bubble with respect to the polarization charge,[6] χ is the surface mobility of the charge carriers on the interface between the media, ρ is the density of the liquid, and σ is the coefficient of surface tension at the liquid-gas interface.

If χ is small and the characteristic time τ_u for the instability to develop is much shorter than the characteristic time τ_q for the electrical potential to equilibrate, i.e., $\tau_u \ll \tau_q$, then the bubble surface can be regarded as nonconducting during the duration of the instability and the charge assumed to be frozen into the surface. When this kind of bubble becomes unstable with respect to the polarization charge that has accumulated on its surface, it breaks up into two equal daughter bubbles that carry equal charges.[7] If, on the other hand, χ is so large that the characteristic time τ_u for instability of the bubble with respect to the surface charge is much longer than the time for equilibration of the electrical potential on the bubble surface, i.e., τ_q, i.e., $\tau_q \ll \tau_u$, then the latter can be assumed to be ideally conducting.

We now estimate the characteristic times τ_q and τ_u for the particular case of a particle with radius $R=10^{-3}$ cm in hexane, assuming that $\rho=0.66\ \mathrm{g/cm^3}$, $\sigma=18.42$ dyn/cm, $\varepsilon=1.88$, the breakdown electric field strength $E=670\ \mathrm{g^{1/2}/cm^{1/2}s}$, and the surface charge carrier mobility is $\chi=1\ \mathrm{cm^{3/2}/g^{1/2}}$. Then the characteristic time for the instability to develop is $\tau_u=0.2\times10^{-4}$ s, the characteristic time for equilibration of the electrical potential on the surface of the bubble is $\tau_q=0.2\times10^{-5}$ s, and $\tau_q \gg \tau_u$. We shall examine this situation in more detail below.

2. The change in the potential energy of a bubble as it stretches out into a spheroid and its volume changes under the influence of the pressure of the electric field $\mathbf{E}$ can easily be written down in a linear approximation with respect to the square of the eccentricity e^2 of the parent bubble,

$$\Delta U = 4\pi\sigma R^2\left(1+\frac{2}{45}e^4\right)-\frac{E^2R^3}{2\varepsilon}\left(1+\frac{2}{5}e^2+\frac{58}{175}e^4\right)$$

$$-PV\ln\frac{V}{V_0}+P_{at}(V-V_0)-4\pi\sigma R_0^2+\frac{E^2R_0^3}{2\varepsilon}, \qquad (1)$$

where V_0 and V are the initial and final volumes of the bubble and P and P_{at} are the pressure of the vapor-gas mixture in the bubble and atmospheric pressure.

Since the equilibrium state of the bubble corresponds to a minimum potential energy, in that state the conditions that the derivatives of Eq. (1) with respect to the independent variables e^2 and R equal zero should be satisfied. This yields two equations in the linear approximation in e^2 which determine the critical conditions for instability of a spheroidal gas bubble in a field $\mathbf{E}$,

$$e^2=\frac{9}{16\pi}w, \quad \frac{(P-P_{at})R}{2\sigma}\equiv\beta=1-\frac{3}{16\pi}w. \qquad (2)$$

Equation (2) shows that, as opposed to a liquid droplet in a uniform electric field $\mathbf{E}$,[6,8] the critical conditions for instability of a gas bubble in $\mathbf{E}$ depend on the gas pressure, which is determined by the dimensionless parameter β.

3. When a bubble elongates in a field $\mathbf{E}$, the charge concentration at the tips of the spheroid can increase so much that, as for a droplet in a field $\mathbf{E}$,[6,8] an instability of high modes of capillary waves develops on the bubble surface and their superposition leads to the formation of emitting protuberances at the tips, from which the emission of charged daughter bubbles begins and carries away the excess polarization charge.[3,6,8] Because of viscous dissipation of the kinetic energy of the daughter bubbles and the braking effect of the electric field of previously emitted bubbles with charges of like sign, the daughter bubbles rapidly come to a halt. Finally, in the neighborhood of the tips of the spheroid at a distance $L=ma$ from the tips of the parent bubble (a is the major semiaxis of the parent bubble and m is a numerical parameter), the daughter bubbles form a cluster.[3] In the following qualitative analysis, the electric field of the cluster of daughter bubbles in the neighborhood of each of the tips is replaced by the field of an equivalent point charge located a distance L from the tip of the parent bubble with a charge equal to the total charge of the entire cluster.

We assume that, as a result of the symmetry of the problem, the n-th emission event from opposite ends of the parent bubble produces two daughter bubbles with charges q_n which are equal in magnitude but of opposite sign ($q_n \ll ER^2$) and radii r_n ($r_n \ll R$). We shall also assume that the n-th daughter bubble, which lies in the total electric field created by the external source, $\mathbf{E}$, plus the polarization charge of the parent bubble, has the shape of an prolate spheroid with eccentricity e_n, which we shall determine by an iteration procedure below.

During breakup of the parent bubble, there is a change in the potential energy of the system which equals the sum of the change in the energy of the surface tension forces, the change in the intrinsic electrostatic energy of the bubbles, the energy of their electrostatic interaction, and the work of compression or expansion of the bubbles owing to nonproportionate changes in the electric and Laplace pressures during emission of daughter bubbles.

We shall assume that the daughter bubbles contain a gas at pressure P_n. Taking the temperature of the system to be constant, it is easy to find the change in the potential energy of the system owing to the n-th breakup. In a linear approximation in terms of small quantities (the radius r_n and charge q_n of a daughter bubble), we obtain

$$\Delta U_n = 8\pi\sigma r_n^2 A(e_n)+q_n^2\frac{B(e_n)}{\varepsilon r_n}+2q_n\frac{ER\Omega(\nu_n)}{\varepsilon}$$

$$+2\frac{q_n}{\varepsilon L}\sum_{i=1}^{n-1}q_i-2P_nV_n\ln\left[\frac{V_n}{V_n^0}\right]+2P_{at}[V_n-V_n^0], \qquad (3)$$

$$A(e_n)=\frac{1}{2}\left((1-e_n^2)^{1/2}+\frac{\sin^{-1}e_n}{e_n}\right)(1-e_n^2)^{-1/6}, \qquad (3a)$$

$$B(e_n)=\frac{(1-e_n^2)^{1/3}}{e_n}\tanh^{-1}e_n, \quad \nu_n=\left(1+\frac{\xi_n}{a^2}\right)^{1/2}, \qquad (3b)$$

$$\Omega(\nu_n)=\frac{e(\nu_n-1)-\nu_n\tanh^{-1}[e(\nu_n-1)(\nu_n-e^2)^{-1}]}{(1-e^2)^{1/3}(\tanh^{-1}e-e)}, \qquad (3c)$$

where ν_n is the distance between the centers of the parent and daughter bubbles, measured in terms of the major semiaxes a of the parent bubble, ξ_n is the spheroidal coordinate of the n-th daughter bubble,[9] and V_n^0 and V_n are the volumes of the daughter bubble before and after the volume changes.

In Eq. (3) the first term describes the surface energy of the two daughter bubbles, the second, their intrinsic electrical energy, the third, the energy of interaction of the daughter bubbles with the parents, the fourth, the energy of the interaction of the cluster of bubbles with the n-th daughter, and the fifth and sixth, the work done by the gas during an isothermal change in the volume of the daughter bubbles.

As in the dispersion of a highly charged bubble,[6,8] the emission process will continue until the Coulomb force which detaches the daughter bubble exceeds the Laplace force $2\pi\sigma r_n^*$ restraining it (r_n^* is the radius of the constriction connecting the n-th detaching bubble to the parent). Given that the field strength at the detachment point is determined by the field of the parent bubble and the field from the cluster of previously detached daughter bubbles, it is easy to find the condition for breakaway of a daughter bubble from the parent,[6,8]

$$\frac{\alpha_n X_n(1-e_n^2)^{1/6}}{8}\leqslant\frac{w}{16}Y_n\left\{T(\nu_n)-\frac{(1-e^2)^{2/3}}{m^2}\sum_{i=1}^{n-1}Y_i\right\}, \qquad (4)$$

where

$$T(\nu_n)=1-\frac{\tanh^{-1}(e\nu_n^{-1})-e\nu_n(\nu_n^2-e^2)^{-1}}{\tanh^{-1}e-e}, \qquad (4a)$$

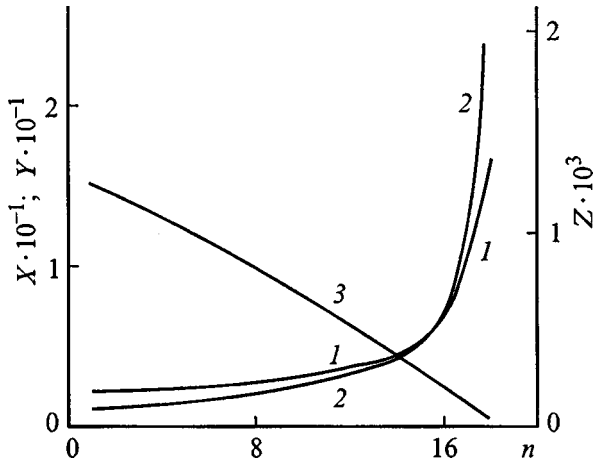

FIG. 1. The dimensionless radii X (*1*), charges Y (*2*), and specific charges Z (*3*) as functions of bubble order number for $w=3.0$ and $m=1$.

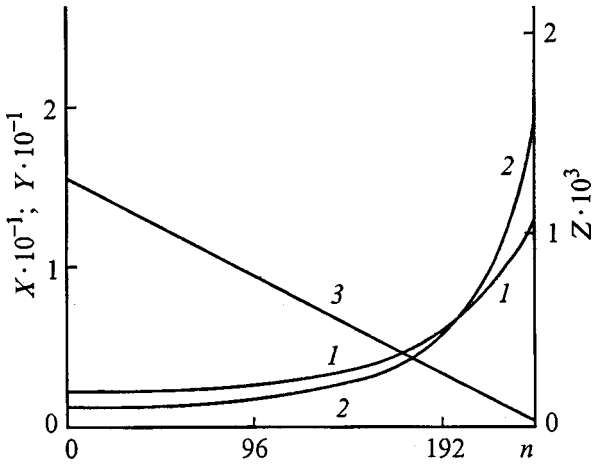

FIG. 3. As in Fig. 1, for $m=10$.

$$Y_n=\frac{q_n}{ER^2},\quad X_n=\frac{r_n}{R},\quad \alpha_n=\frac{r_n^*}{b_n},\tag{4b}$$

Y_n and X_n are the dimensionless charge and radius of the n-th daughter bubble, and b_n is the minor semiaxis of the n-th daughter bubble.

The first term in the curly brackets of Eq. (4) is characterized by the electric field strength created by the parent bubble at the point where the n-th daughter bubble breaks away and the second is determined by the electric field strength of the cluster of previously emitted daughter bubbles.

Since the process of ejecting excess charge by emitting a large number of daughter bubbles is a nonequilibrium second order, stationary process, we require, in view of the Onsager minimum energy dissipation principle for nonequilibrium processes,[10,11] that the change in the potential energy of the system during emission of the next bubble should be extremal, i.e., that the condition $\partial(\Delta U_n)/\partial q_n=0$, $\partial(\Delta U_n)/\partial r_n=0$ be satisfied. Taking the partial derivatives of Eq. (3), we obtain two equations besides Eq. (4) for finding the three unknowns X_n, Y_n, and ν_n:

$$Y_n\frac{B(e_n)}{X_n}+\Omega(\nu_n)+\frac{(1-e^2)^{1/3}}{m}\sum_{i=1}^{n-1}Y_i=0,\tag{5}$$

and

$$X_nA(e_n)-\frac{w}{16\pi}Y_n^2\frac{B(e_n)}{X_n^2}-\eta_n\beta X_n^2=0,\tag{6}$$

where

$$\eta_n=\frac{P_n-P_{\text{at}}}{P-P_{\text{at}}}.\tag{7}$$

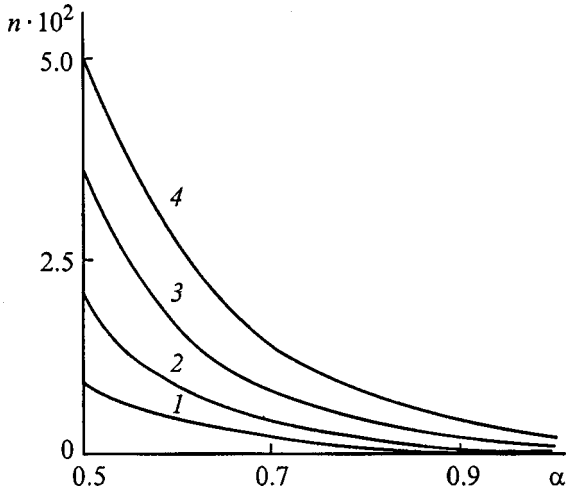

FIG. 2. The number of daughter bubbles as a function of the undefined parameter for $w=2.0$ (*1*), 2.5 (*2*), 3.0 (*3*), 3.5 (*4*).

4. For $e^2=0.7$, $\eta_n=\eta=0.6$ (calculations show that the daughter bubble has a charge close to the limit in the Rayleigh stability sense, so it expands), $\beta=0.75$, and various fixed values of the parameters w, α_n, and m, the system of Eqs. (4)–(6) can be used to calculate the dimensionless radii and charges of the daughter bubbles. In these calculations (illustrated in Fig. 1), it became clear that, as for the instability of a highly charged droplet,[6,8] the radii and charges of the daughter bubbles increase with the number of the daughter. The α_n were assumed equal to $\alpha=0.9$, by analogy with a droplet, since calculations of the breakup of charged droplets[8] with $\alpha=0.9$ give the best agreement between the calculations and experimental data.[8,12] $m=1$ was chosen on the basis of photographs.[3]

The calculations also showed that, as for a droplet, an increase in the Taylor parameter w of the parent bubble leads to a rise in the number of daughter bubbles,[13] as can be seen in Fig. 2, where the number n of emitted daughter bubbles is plotted as a function of the undetermined parameter α for different fixed values of the Taylor parameter w of the parent bubble.

The number of daughter bubbles rises, both when the Taylor parameter is increased and when the distance to the clusters of daughter bubbles becomes greater. (A situation with $L\gg a$ may occur, for example, for bubbles in low viscosity liquids.) As L increases, the number of daughter

bubbles increases, while their radii and charges change very little, as can be seen in Fig. 3 (which illustrates the calculations for $m=10$). It is easy to see from a comparison of Figs. 1 and 3 that, as L increases from $L=a$ to $L=10a$, the number of daughter bubbles increases by an order of magnitude.

CONCLUSION

When a bubble in a liquid dielectric in an external uniform electrostatic field becomes unstable, it emits several dozen fine, highly charged daughter bubbles which form two clusters at each of the emission protuberances in the neighborhood of the tips of the parent bubble. Increasing the distance from the surface of the parent bubble to the clusters of previously emitted bubbles by an order of magnitude (by reducing the viscosity of the liquid) leads to an increase in the number of daughter bubbles by an order of magnitude, as well. Increasing the Taylor parameter for the parent bubble leads to a substantial increase in the number of emitted daughter bubbles.

[1] V. V. Glazkov, O. A. Sinkevich, and P. V. Smirnov, Teplofiz. Vys. Temp. **29**, 1095 (1991).
[2] I. V. Pilaeva, O. A. Sinkevich, and P. V. Smirnov, Teplofiz. Vys. Temp. **30**, 367 (1992).
[3] G. G. Garton and Z. Krasucki, Trans. Faraday Soc. **60**, 211 (1964).
[4] S. O. Shiryaeva and A. I. Grigor'ev, Zh. Tekh. Fiz. **64**(3), 13 (1994) [Tech. Phys. **39**, 229 (1994)].
[5] S. O. Shiryaeva and A. I. Grigor'ev, J. Electrost. **34**, 51 (1995).
[6] A. I. Grigor'ev and S. O. Shiryaeva, J. Phys. D **23**, 1361 (1990).
[7] A. I. Grigor'ev, V. A. Koromyslov, and A. N. Zharov, Pis'ma Zh. Tekh. Fiz. **23**(19), 60 (1997) [Tech. Phys. Lett. **23**(10), 760 (1997)].
[8] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61**(3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].
[9] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* [in Russian], Nauka, Moscow (1982), 620 pp.
[10] S. O. Shiryaeva and A. I. Grigor'ev, Zh. Tekh. Fiz. **65**(2), 11 (1995) [Tech. Phys. **40**, 122 (1995)].
[11] C. T. Yang C.T. and C. C. S. Song, *Encyclopedia of Fluid Mechanics. Flow Phenomena and Measurement*, Houston Book Division, Gulf Publishing Company, London–Paris–Tokyo (1986), Vol. 1, pp. 353-400.
[12] A. I. Grigor'ev and S. O. Shiryaeva, Izv. Ross. Akad. Nauk Mekh. Zhidk. Gazov, No. 3, pp. 3–32 (1994).
[13] S. O. Shiryaeva and A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **19**(18), 87 (1993) [Tech. Phys. Lett. **19**, 600 (1993)].

Translated by D. H. McNeill

Microwave streamer discharge in a supersonic air flow

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

Moscow Radio Engineering Institute, Russian Academy of Sciences, 113519 Moscow, Russia
(Submitted February 9, 1998; resubmitted January 30, 1999)
Zh. Tekh. Fiz. **69**, 14–18 (November 1999)

Data are presented from experiments on the ignition of a pulsed, triggered microwave streamer discharge at the focus of a cm-band *TEM* wave in an immersed supersonic air jet. It is shown experimentally that for velocities of the air in the jet up to 500 m/s, the structure of the discharge remains qualitatively unchanged and retains its streamer character. The finite size of the transverse cross section of the jet determines some features of the discharge. © *1999 American Institute of Physics.* [S1063-7842(99)00311-6]

Pulsed microwave gas discharges in electromagnetic beams with a *TEM* field structure at centimeter wavelengths λ in gases at pressures p above a threshold p_{thr} determined by the gas species, exist in the form of streamers.[1] In integrated photographs with exposure times t_e exceeding the duration t_p of the microwave pulse, this kind of discharge shows up as a complicated composite of thin plasma channels. Studies of the time evolution of these discharges show that the channels, or streamers, that make up the discharge are continually lengthening and branching.

One of the main properties of a developed microwave streamer discharge is that it absorbs almost all the electromagnetic energy incident on it. This is a significant difference between microwave streamer discharges and spatially uniform discharges.

Another important feature of microwave streamer discharges is that the streamers of which they are comprised, like streamers in a dc field, can grow into a region of the electromagnetic beam where the amplitude of the electric component of the initial field, E_0, is substantially lower than the critical breakdown amplitude E_k (Ref. 2). In practice, this makes it possible to produce microwave streamer discharges in an electromagnetic beam for which $E_0 \ll E_k$ throughout its entire volume. Here the conditions for breakdown are created by special measures only within its local region. In integrated photographs, such subcritical triggered discharges show up as a ''tangle'' of plasma channels which expand with t_p from the initiation site to the discharge boundary facing the microwave source; that is, the discharge propagates toward the radiator and one can introduce the concept of a discharge front velocity V_{fr}. High speed photography of subcritical discharges shows that the main energy transfer processes in the discharge region take place right at its front.[3] In the front, isolated streamer segments comparable to $\lambda/2$ successively form resonant plasma dipoles that ensure absorption of the electromagnetic energy. Behind the ''front,'' the discharge plasma actually begins to decay.

A microwave streamer discharge at $\lambda = 8.5$ cm was produced in air with $p > p_{\text{thr}} = 60$ Torr and a triggered variant with a developed structure was produced with $E_0 > 1.5$ kV/cm and $t_p > 10\ \mu$s.[4] Scanning photographs show that the resonant portions of the streamers which appear in the discharge front radiate intensely in the visible spectral range for a time t_{sv} of several microseconds.[3] Under these conditions, the average velocity of the front can be estimated from the empirical formula[5]

$$V_{\text{fr}} = (1/\sqrt{3})(3\times 10^4 \sqrt{p} + 1.7\times 10^{-2} E_0 p\lambda); \quad \text{cm/s}, \qquad (1)$$

where p has dimensions of Torr, E_0 is in V/cm, and λ is in centimeters. The second cofactor is the average rate of growth of the streamers which form the discharge, V_{str}, and the first reflects the fact that they propagate predominantly upstream and perpendicular to the microwaves.

It is clear from Eq. (1) that for a pressure p of hundreds of Torr, V_{str} is of the order of 10^6 cm/s. For comparison, we also introduce the value of E_k in air:[2]

$$E_k = 1.2\times 10^{-15} n; \quad \text{V/cm}, \qquad (2)$$

where n is the density of air molecules in cm^{-3}.

The two properties of this type of discharge noted above are decisive to its practical applications. Indeed, on one hand, a microwave streamer discharge uses the electromagnetic energy incident on it with high efficiency, while, on the other, its triggered subcritical variant requires comparatively low power and practically accessible microwave sources, even at very high gas pressures. At the same time, for certain applications, such as in aerodynamics,[6–9] it is necessary to know how this kind of discharge behaves in gas flows, including supersonic ones. Using Eq. (1) as a reference, we may assume that, for flow velocities $V_F \leqslant 10^5$ cm/s, the discharge structure and properties will not change significantly. Nevertheless, this assumption requires experimental verification.

In this paper we present data from experiments on the ignition of pulsed, triggered subcritical microwave streamer discharges in a *TEM* beam in a supersonic air jet. Since microwave streamer discharges produced in electromagnetic beams in gas flows have been little studied, at this stage of our study primary attention has been directed at the fundamental aspects of the problem. The problem was: will a subcritical microwave discharge at $p > p_{\text{thr}}$ develop at all in a flow and retain its streamer structure?

In the experiments we used an apparatus that has been described previously,[1,3,4] with some modifications. The ex-

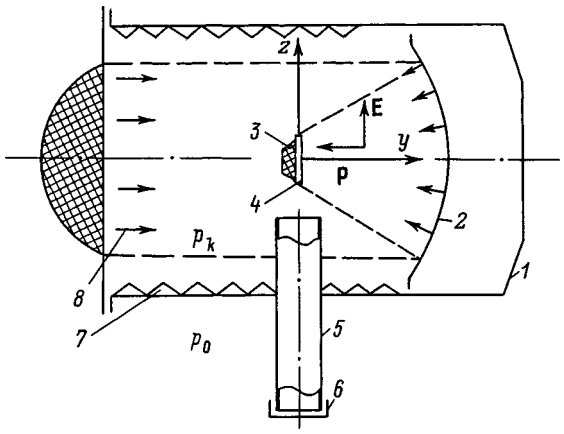

FIG. 1. A sketch of the experimental apparatus for igniting subcritical triggered microwave streamer discharges in supersonic air flows: *1* — leakproof vessel, *2* — focusing mirror, *3* —ruler, *4* — trigger dipole, *5* — tube, *6* — breakable diaphragm, *7* — microwave absorbers, *8* — *TEM* wave.

perimental arrangement is shown in Fig. 1. Shown there is the vacuum chamber (*1*), a mirror (*2*) for focusing the microwave radiation, a triggering microwave dipole (*4*) located at the focus, and a dielectric tube (*5*) which shapes the supersonic jet that flows around the trigger. In the experiments, the microwaves were focused onto the center of the vessel. The vessel and tube were first pumped to a pressure p_k. At a given time the outer end of the tube was opened to the atmosphere ($p_0 = 760$ Torr) and air began to flow into the vessel. Then the microwave generator was turned on after a fixed delay. The discharge that developed in the focal region was photographed with $t_e \geqslant t_p$.

Linearly polarized *TEM* radiation with a plane phase front was used in the experiments. It was incident on the vessel at the focusing mirror. In the focal region the transverse distribution of the field amplitude of the beam was close to Gaussian with a characteristic size of 5 cm. Along the beam axis, the field was roughly uniform over a length of a few centimeters. The microwaves had a wavelength of $\lambda = 8.5$ cm, a field amplitude of $E_0 \leqslant 7.2$ kV/cm at the focus, and a pulse duration of $t_p = 40\ \mu$s.

The trigger consisted of a copper wire with a diameter of 0.8 mm and a length of 4 cm. It was placed at the focus of the electromagnetic beam parallel to $\mathbf{E}_0$ and attached to the middle of a plexiglas ruler (item *3* in Fig. 1) with a width of 25 mm and a thickness of 6 mm. The ruler passed through the focus and was perpendicular to the direction of the radiation, which was characterized by a Poynting vector $\mathbf{P}_0$ and field vector $\mathbf{E}_0$. Its wide side was parallel to $\mathbf{E}_0$. In order to reduce the effect of the ruler on the air flow in the region of the proposed discharge, the ruler was chamfered in its transverse cross section, as shown in Fig. 1. The trigger was fastened to the flat side of the ruler, which was turned toward $\mathbf{P}_0$.

The dielectric tube used to shape the supersonic jet had an inner diameter $d_{ssj} = 5.4$ cm and a length of 81 cm. The axis of the tube was collinear with the axis of the trigger. The open outlet end of the tube was located 10 cm form the focus. The outer end of the tube was sealed from the atmosphere by a diaphragm of polyethylene- terephthalate, a dielectric film with a thickness of 20 μm (item *6* in Fig. 1). A nichrome wire with a diameter of 0.1 mm and a length of 3 cm was placed next to the outside of the diaphragm. The wire was connected to the discharge circuit of a high voltage capacitor. When the capacitor was discharged, the wire exploded and broke the film. From that time, a linear influx of air into the vessel began at an experimentally measured rate of 0.38 Torr/ms.

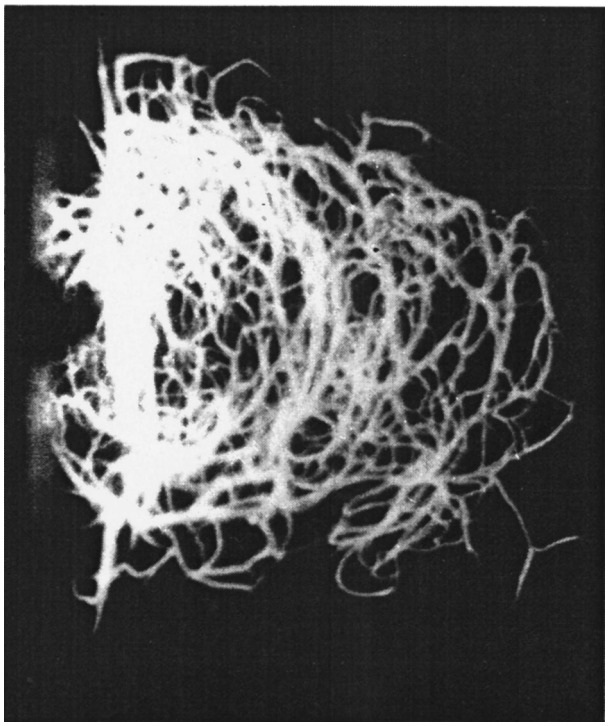

FIG. 2.

In the experiments the time Δt between the breaking of the film and the delivery of the microwave pulse was specified. The value of Δt was estimated in advance from the expected flow velocity V_F. Here it was kept in mind that, in the focal region, the flow had already been able to form but there was, as yet, no reflection of the jet from the vessel wall opposite the tube, located 45 mm from the focus. During the experiments the value of Δt was determined more accurately, since it was difficult to predict the time the diaphragm would break. (In the experiments it was found that this time lay within a range of a few milliseconds.)

Figures 2–7 show pictures of triggered subcritical microwave streamer discharges obtained during the experiments. In them the radiation travels from right to left and $\mathbf{E}_0$ lies in the plane of the figure and perpendicular to $\mathbf{P}_0$. In order to explain the perspective in the pictures we introduce a right-hand Cartesian coordinate system with its origin at the focus of the electromagnetic beam, its *Y* axis directed opposite $\mathbf{P}_0$, and its *Z* axis directed upward in the figures. Now, if we superimpose spherical coordinates on it, the axis of the camera objective forms a latitudinal angle $\theta = 110°$ relative to the *Z* axis and a longitudinal angle $\varphi = 20°$ with respect to the *X* axis. In the photographs the length of the trigger dipole can serve as a scale length for the image.

Figure 2 shows a discharge without a flow at

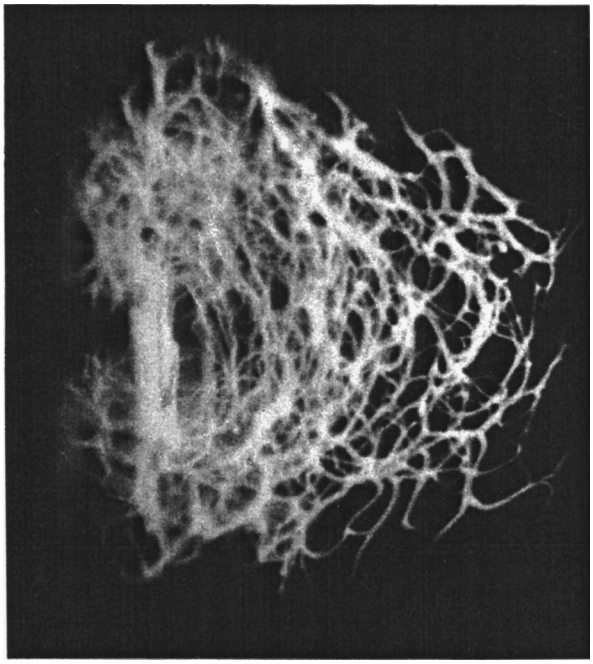

FIG. 3.

$p_k = 300$ Torr and $E_0 = 7.2$ kV/cm. The characteristic ''tangle'' of streamer channels can be seen and, to the left, the vertical luminosity broken in the middle is a reflection from the vertical piece of the mirror clamp, which is far from the discharge region (about 35 cm along the X axis). The bright uniform formation on the left of the ''tangle'' is a discharge over the surface of the ruler at the place where it comes into contact with the trigger dipole. Above and below it are the more brightly emitting ends of the trigger dipole (the luminosity of the lower end is somewhat more distinct). The discharge extends about 6 cm upstream of the micro-

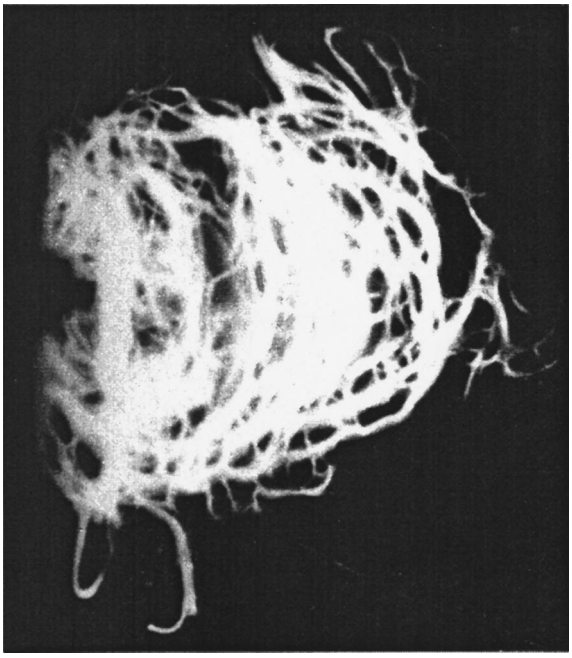

FIG. 4.

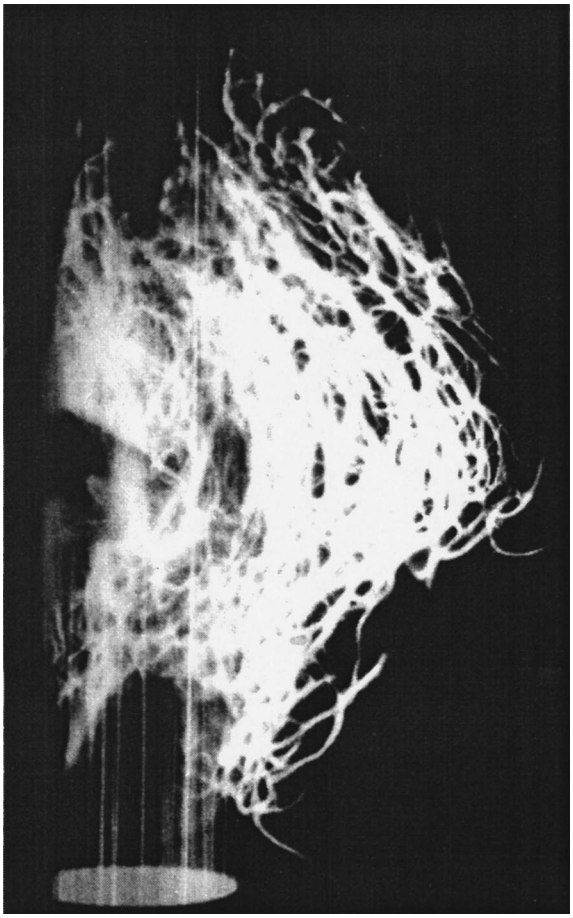

FIG. 5.

wave radiation, which corresponds to $V_{\rm fr} \cong 1.5 \times 10^5$ cm/s and, therefore, to $V_{\rm str} \cong 2.5 \times 10^5$ cm/s. (The $V_{\rm fr}$ estimated from the photograph can be compared with the value calculated using Eq. (1). Here it should be kept in mind that Eq. (1) represents averaging both over t_p and over a series of pulses; that is, the average estimate of $V_{\rm fr}$ from Fig. 2 corresponds to just this situation.)

Figure 3 shows a discharge with the same $p_k = 300$ Torr and $E_0 = 7.2$ kV/cm, but with a jet, for $\Delta t = 10$ ms. The outlet end of the tube that forms the jet is visible at the bottom of the picture. It is illuminated by the spark from the exploding wire which broke the seal on the tube. Tracks formed by ''fragments'' of the wire and diaphragm entrained by the air flow can also be seen. To some extent, they provide a visual representation of the jet. As before, reflections and illumination of the ruler by the discharge can be seen to the left. It is evident that the distance moved by the discharge toward the radiator when the jet was present did not change significantly; that is, $V_{\rm fr}$ and $V_{\rm str}$ are essentially unchanged. In this geometry, in the initial stages of its development the discharge is in the flow and then propagates in air that has not yet moved. This shows up in the photograph as streamer channels that are somehow ''smeared out'' along the flow when they lie within the flow (cf. Fig. 2) but have distinct profiles outside the flow.

Figure 4 shows a discharge without a flow and with $p_k = 100$ Torr and $E_0 = 7.2$ kV/cm. At this p_k the streamer chan-

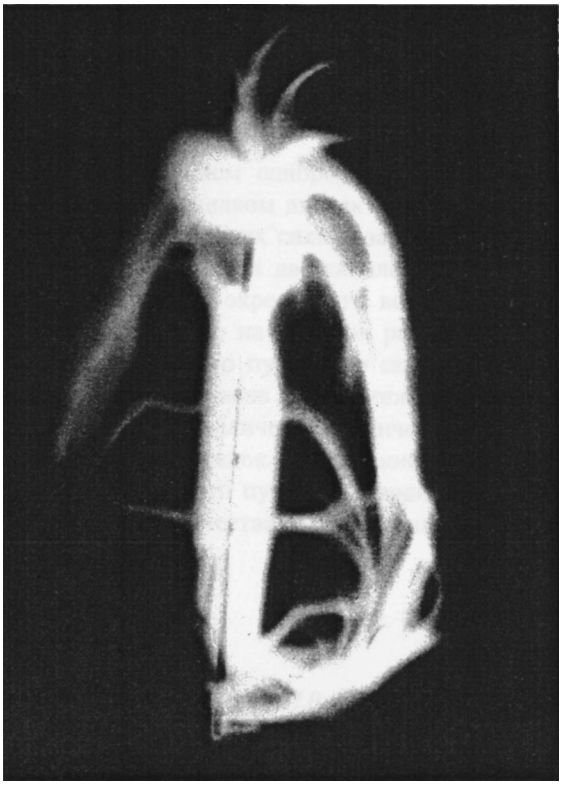

FIG. 6.

FIG. 7.

nels which form the discharge are thicker and more diffuse than in Fig. 2. Figure 5 shows a discharge with the same p_k and E_0 but with a jet at $\Delta t = 50$ ms. It shows a typical streamer discharge. At the same time, the limited size of the transverse cross section of the jet gives it a certain aspect. The discharge mainly develops around the jet, as if it surrounded the jet and made its boundary visible.

In order to examine the discharge inside the jet in more detail, Figs. 6 and 7 show discharges without and with a jet, respectively, for $p_k = 100$ Torr and $\Delta t = 10$ ms, but already with a lower $E_0 = 1.7$ kV/cm. These photographs were obtained with a longer focal length lens, i.e., with a shorter depth of field. The trigger is clearly visible in these pictures.

In Fig. 6, as before, the illumination of the ruler by the discharge is clearly visible in the central portion of the trigger dipole along its surface at the place where the dipole comes into contact with it. Near its upper end, vaporization of the dipole material can be observed. Vaporization takes place at the point where its surface comes into contact with the high temperature plasma channels.

In Fig. 7, as before, the boundaries of the flow can be determined from the tracks. Here it is apparent that the picture of the discharge has changed with addition of a jet. Under these conditions there is no discharge along the dipole at the ruler surface. The discharge stretches out significantly along the flow. It occupies a larger volume. At the same time, the discharge was initiated by the dipole, which lies within the flow. The lower end of the dipole is subjected to vaporization, i.e., the flow has not reduced the temperature of the plasma channels significantly. The channels emerging from the upper end of the dipole were substantially ''smeared out'' along the flow.

Let us comment on these results, given the estimates of the characteristic parameters of the jet in Figs. 3, 5, and 7. We shall assume that over the entire length of the tube there is a critical flow[10] (in the experiments, always, $p_0/p_k > 1.9$). Then, regardless of p_k, at its outlet cross section we obtain a static pressure $p_{ssj} = 400$ Torr, air temperature $T_{ssj} = 250$ K, and air flow velocity $V_{ssj} = 310$ m/s, for a molecular density of $n_{ssj} = 1.9 \times 10^{19}$ cm^{-3}.

When $p_{ssj} > p_k$, as it comes out of the tube the jet takes the form of Mach ''rolls.'' First it expands, while the static pressure p_F in the flow decreases. For $p_k = 300$ Torr the pressure p_F equilibrates with p_k as the jet diameter d_F increases beyond its outlet size d_{ssj} by only 2%. At this cross section of the jet, the Mach number of the flow will be $M_F = 1.23$, with $V_F = 370$ m/s, while the density of molecules in the flow will be $n_F = 1.3 \cdot n_k$, where $n_k = 10^{19}$ cm^{-3}. For $p_k = 100$ Torr, at the jet cross section where p_F equals p_k these quantities are $d_F = 1.3\, d_{ssj}$, $V_F = 500$ m/s, $n_F = 1.8 n_k$, and $n_k = 3.4 \times 10^{18}$ cm^{-3}.

These numbers show that for $p_k = 300$ Torr the density n_F at this cross section is 30% higher than n_k; that is, in the flow E_k is 30% higher than the critical field in the motionless air surrounding the jet. According to Eq. (2), the latter is $E_k = 12$ kV/cm. In this case, therefore, both in the flow and outside it, only a subcritical discharge can develop with $E_0 = 7.2$ kV/cm. Thus, the discharge shown in Fig. 3 is subcriti-

cal and triggered, and develops in the flow in its initial stage. The diffuse smearing of the streamer channels inside the flow seen in the photographs does not necessarily reflect a change in the growth of the streamers and in their ability to dissipate the electromagnetic energy. It only characterizes the duration of their luminosity. Note, however, that for the time t_{sv} estimated in Ref. 3, the absolute broadening along the flow of streamer channels perpendicular to $\mathbf{V}_F$ should be about a millimeter. In Fig. 3 the ''broadening'' is substantially greater. Further study is required to explain this disagreement.

For $p_k = 100$ Torr, $E_k = 4$ kV/cm in the vessel and 7.2 kV/cm in the flow. Evidently, in this case, for $E_0 = 7.2$ kV/cm the discharge should mostly ''surround'' the jet, as shown in Fig. 5. For $E_0 = 1.7$ kV/cm, which corresponds to Fig. 7, the conditions are subcritical everywhere, so the discharge can only be triggered within the jet and outside it. The major differences between the discharge of Fig. 7 and the pattern shown in Fig. 6 are most likely associated with an increased molecular density in the region of the trigger dipole when the flow is present.

These experiments on the ignition of microwave streamer discharges in supersonic gas jets with velocities of up to 500 m/s have shown that discharges under these conditions are still streamer discharges. This result offers a realistic possibility that this type of discharge might be used in applied aerodynamics for example, by transmitting electromagnetic energy from aircraft[8,9] in order to deal with the shock waves from supersonic flight or to change the aerodynamic characteristics of the aircraft. At the same time, in order to study microwave discharges in jets, as for experiments in a dc field,[6] the finite transverse dimensions of the jet and the difference between the parameters of the gas in the jet and in the surrounding space must be taken into account. Neglect of these factors can lead to significant misinterpretations of the observed phenomena.

We thank K. V. Aleksandrov for help in doing the experiments.

[1] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **55**, 389 (1985) [Sov. Phys. Tech. Phys. **30**, 228 (1985)].
[2] A. D. MacDonald, *Microwave Breakdown in Gases* [Wiley, New York (1966); Mir, Moscow (1969), 205 pp.].
[3] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Pis'ma Zh. Tekh. Fiz. **18**(22), 34 (1992) [Sov. Tech. Phys. Lett. **22**(11), 737 (1992)].
[4] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **64**(1), 74 (1994) [Tech. Phys. **39**, 40 (1994)].
[5] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **65**(5), 21 (1995) [Tech. Phys. **40**, 416 (1995)]
[6] L. P. Grachev, N. N. Gritsov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **61**(9), 185 (1991) [Sov. Phys. Tech. Phys. **36**, 1073 (1991)].
[7] V .V. Vitkovskiĭ, L. P. Grachev, N. N. Gritsov *et al.*, Teplofiz. Vys. Temp. **28**, 1156 (1990).
[8] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **55**, 972 (1985) [Sov. Phys. Tech. Phys. **30**, 586 (1985)].
[9] I. I. Esakov, Izv. Sib. Otdel. Akad. Nauk SSSR Ser. Tekh. Nauk, No. 1, 98 (1990).
[10] L. G. Loĭtsyanskiĭ, *Mechanics of Liquids and Gases* [in Russian], Nauka, Moscow (1979), 904 pp.

Translated by D. H. McNeill

Range of existence of self-sustained, subcritical microwave streamer discharges

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

Moscow Radiotechnical Institute, Russian Academy of Sciences, 113519 Moscow, Russia
(Submitted February 9, 1998; resubmitted January 30, 1999)
Zh. Tekh. Fiz. **69**, 19–24 (November 1999)

Data are presented from a study of subcritical microwave streamer discharges in air in an electromagnetic beam. It is shown that, depending on the degree of subcriticality, this kind of discharge can have two forms: a self-sustained discharge and a discharge that is attached to the initiating trigger. The range of subcriticality for the initial field is determined as a function of air pressure within which the self-sustained developed discharge form exists. © *1999 American Institute of Physics.* [S1063-7842(99)00411-0]

It has been proposed[1] that subcritical microwave discharges be used to clean the earth's atmosphere of the Freons which are destroying its ozone layer. It is assumed that a discharge will be created in an electromagnetic beam formed by an earthbound antenna. The discharge will be initiated at the specified height above the earth's surface by a system of triggering dipoles and then maintained continuously by the electromagnetic radiation. Using an electromagnetic beam with fields significantly below the critical field for self-sustained breakdown of air at a given pressure makes it possible to reduce the microwave power emitted by the antenna to a practically realizable magnitude. Employing a streamer discharge offers the possibility of using just a small number of initial triggering dipoles. Here we are using the tendency of these discharges to break away from the trigger and then develop continuously and in a self-sustained fashion in a subcritical field as growing, branching streamer channels. At the same time, as was noted in Ref. 1, the minimum external field level at which a discharge will break away from the trigger and develop independently is unknown at present. It is very important to know this. It may turn out, for example, that it places a requirement on the microwave power required to implement this method of decontaminating the atmosphere.

In this paper we present some results from experiments to determine the lower limit on the field for existence of a self-sustained (independently developing) subcritical microwave discharge as a function of the air pressure.

The experiments were done on an apparatus which has been described in detail elsewhere.[1] Linearly polarized electromagnetic radiation with a wavelength $\lambda = 8.5$ cm and pulse duration $\tau = 40\ \mu$s was focused onto the center of a vacuum vessel. The characteristic size of the focal spot was a few centimeters. The initial amplitude of the electric field at the focus was varied from $E_0 = 0.95$ to 2 kV/cm and determined to within 10%. The experiments were done in a single microwave pulse regime. The pauses between pulses were at least 1 min. The air pressure p in the vessel could be varied from tens of Torr to atmospheric and was measured with an accuracy of ± 1.5 Torr. A triggering dipole was placed at the focus parallel to $\mathbf{E}_0$. It consisted of a copper wire with a diameter of $2a = 0.8$ mm and a length $2L$ of 2 or 3 cm.

In the experiments the discharge region was photographed with an exposure time $t_e \gg \tau$. In the photographs shown below the microwave radiation moves from right to left and $\mathbf{E}_0$ is vertical. The dimensions of the triggers served as a scale for the images.

The experimental results for each of the trigger dipole lengths indicated above are shown in Figs. 1 and 2 as a matrix of photographs in which the columns correspond to a particular pressure p and the rows, to a definite amplitude E_0 of the electromagnetic field. The photographs shown in Fig. 1 correspond to $2L = 2$ cm, $p = 60$, 100, 150, 200, and 400 Torr and $E_0 = 0.95$, 1.2, 1.4, 1.6, 1.8 and 2 kV/cm. The maximum pressure p is limited by the feasibility of breakdown in air with a given trigger at the maximum E_0 used in the experiments. (For $E_0 = 2$ kV/cm and $p > 400$ Torr, the air did not break down with this trigger.) The photographs shown in Fig. 2 were taken for $2L = 3$ cm. In this matrix, besides columns and rows analogous to Fig. 1, there is a column for $p = 760$ Torr.

For analyzing the experimental data it is convenient to introduce the concept of the degree of subcriticality of the initial field, $\Psi = E_k/E_0$, where the critical breakdown field $E_k = 40p$ in V/cm if p is given in Torr. It is evident from Figs. 1 and 2 that the shape of the discharge differs substantially for low and high subcriticalities. For low Ψ the discharge has the extensively described streamer discharge configuration.[2,3] As an example, Fig. 3 shows a photograph of this sort of discharge, on a larger scale than in Fig. 2, for $p = 150$ Torr and $E_0 = 2$ kV/cm, i.e., $\Psi = 3$. This is a typical tangle of plasma channels. It extends roughly 5 cm toward the microwave radiation, which corresponds to an average propagation velocity for the discharge front of $V_{fr} = 1.25 \times 10^5$ cm/s or an average rate of growth of the constituent streamer channels of $V_{str} = 2.2 \times 10^5$ cm/s.[4]

At high Ψ the discharge character changes significantly. As an example, Fig. 4 shows a photograph, on a larger scale than in Fig. 2, corresponding to $p = 150$ Torr and $E_0 = 0.95$ kV/cm, i.e., $\Psi = 6.3$. There it can be seen that streamer channels move out from the poles of the trigger dipole, but throughout the entire microwave pulse they continue as if attached to it. Outwardly, this is a completely different type of microwave discharge. In the following we

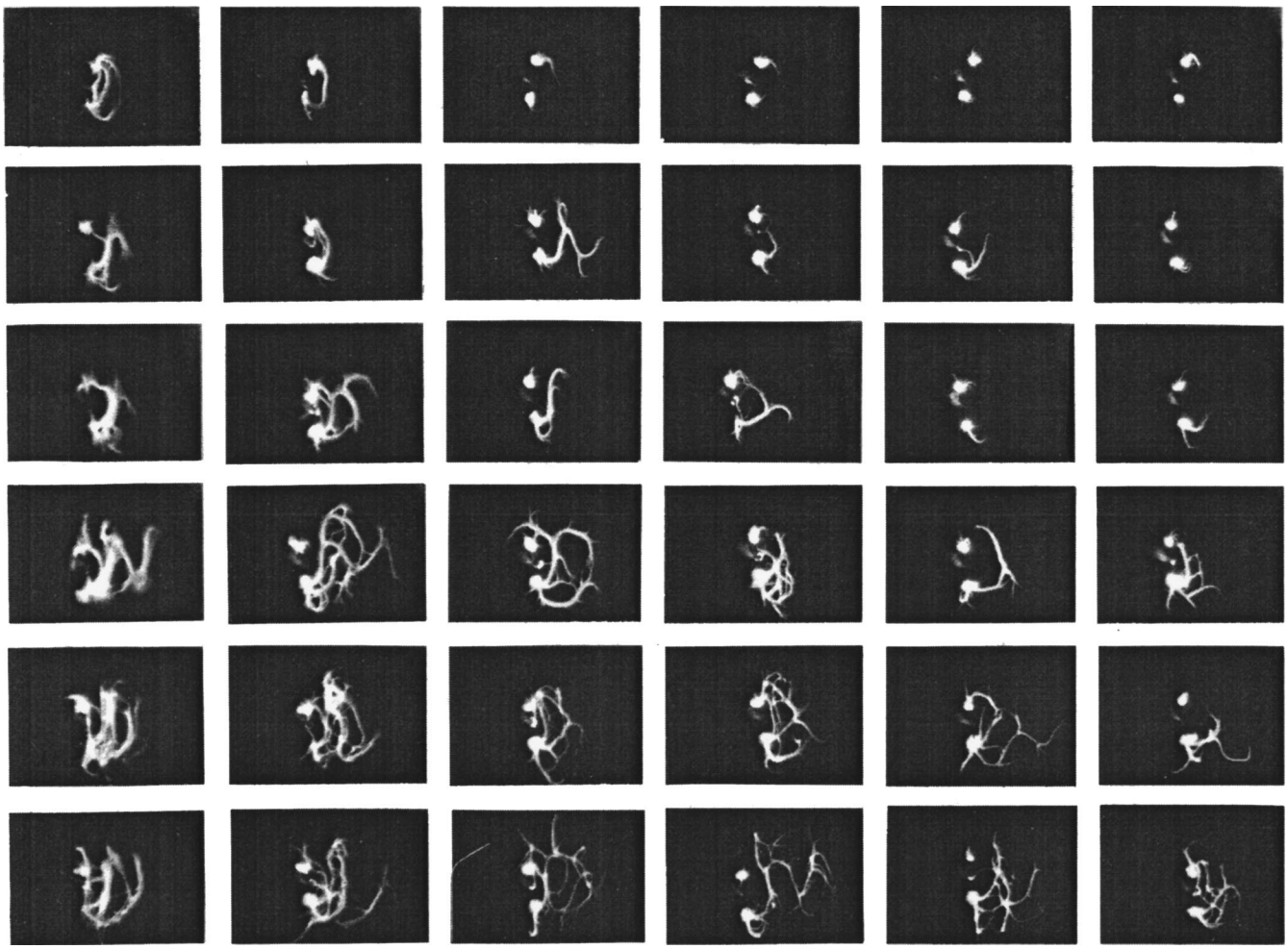

FIG. 1. External appearance of triggered subcritical microwave discharges at different air pressures and field amplitudes: $2L=2$ cm, $2a=8\times10^{-2}$ cm.

shall refer to it as an attached subcritical microwave streamer discharge.

The limits with respect to E_0 or Ψ for the transition from an independently (self) sustained subcritical discharge to an attached subcritical discharge can be determined from Figs. 1 and 2. First, let us formulate the criterion by which we shall be guided in determining this boundary. Here the main difficulty is in the limited capabilities of the experimental apparatus: the short duration of the microwave pulse and the small volume of the focal region with a comparatively uniform field. Under these experimental conditions, even a self-sustained discharge cannot move very far from the trigger. In addition, as Figs. 1 and 2 show, as E_0 is lowered, the ''density'' of the streamer channels decreases owing to a reduction in the size of the electromagnetic beam region within which this type of discharge can develop. These difficulties are enhanced by the tendency of the microwave streamer discharge, itself, to form a complicated spatial configuration which is not repeated in different microwave pulses. Thus, photographing the discharge region from one aspect means that it is not always possible to determine the shape and length of each of the streamers, even if there is a small number of them. (Producing a large number of shots, on the other hand, makes the experiment extremely time consuming.)

In order to develop a criterion, we used the data of Ref. 2. Figure 5 shows photographs, taken from that reference, of a discharge initiated by a small sphere with a diameter of $2a=2.5$ mm at a low pressure $p=20$ Torr, $E_0=0.8$ kV/cm, and $\tau=3$ (Fig. 5a) and 20 μs (Fig. 5b). This is a typical diffuse microwave discharge. Beginning at the poles of the sphere, it expands along $\mathbf{E}_0$ as a result of a field-ionization process to a size no greater than $\lambda/2$ and then forms a diffuse plasma loop directed toward the microwave source. Figure 5a shows that the time to form this structure is no more than a few microseconds. For the rest of the time before the end of the microwave pulse, the size of the plasmoid increases (Fig. 5b), but the velocity of this process is incomparably lower than the characteristic velocity at which the initial plasma loop is formed. As a whole, this discharge was essentially attached to the trigger throughout the time τ. (The smooth increase in its size is caused by an ionization thermal process by which the boundary of a diffuse discharge moves toward the microwave source and is not considered here.)

Figures 1 and 2 show that within a certain range of Ψ, a streamer discharge can also form a loop that is attached to the poles of the trigger and is turned toward the microwave source. This loop can consist of one or several streamer channels as shown, for example, in Fig. 6, which shows the

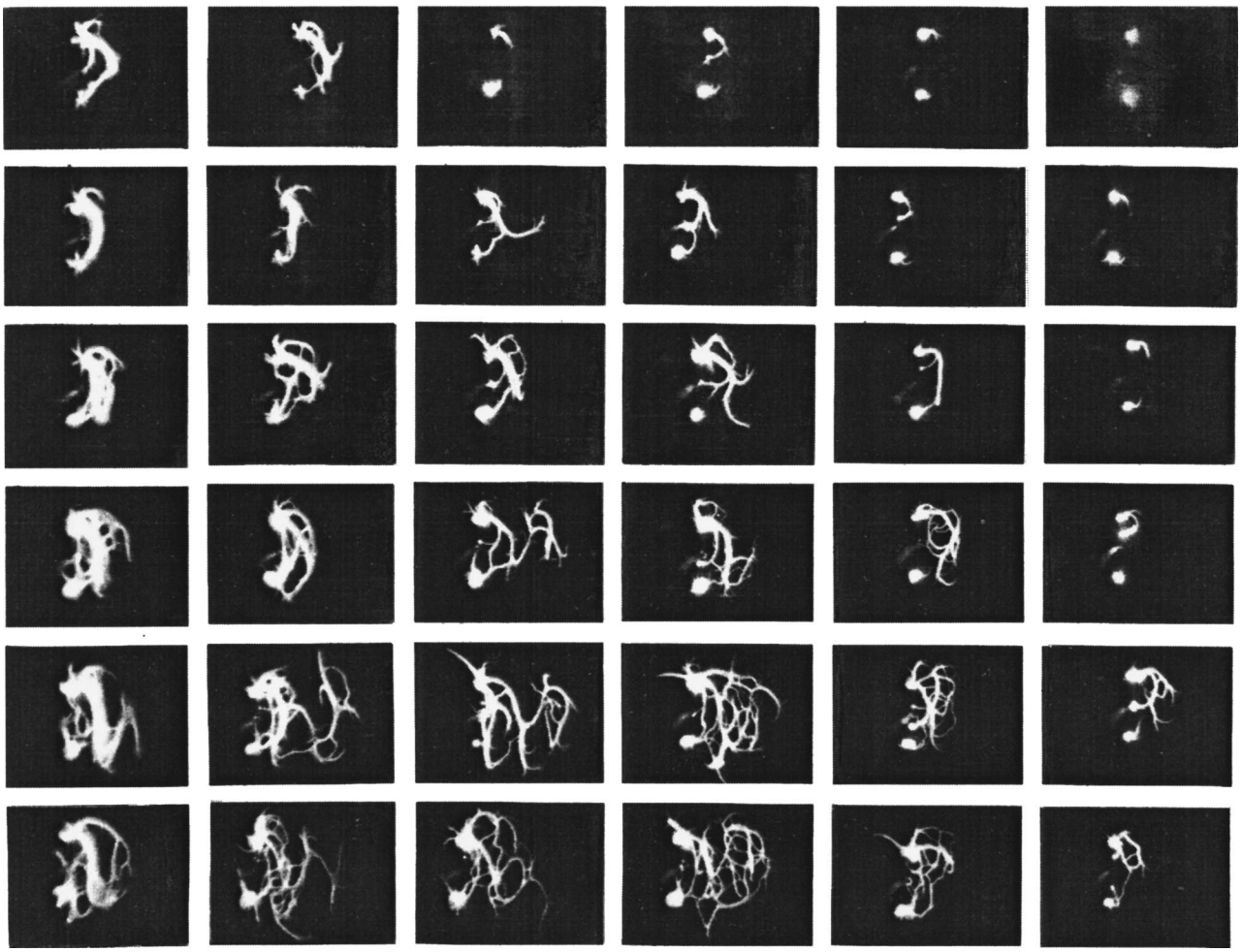

FIG. 2. As in Fig. 1, for $2L=3$ cm.

element of the matrix of photographs for $2L=2$ cm with $p=100$ Torr and $E_0=1.2$ kV/cm.

By analogy with the diffuse form of a discharge shown in Fig. 5, we shall regard a streamer subcritical discharge as ''attached'' if it only forms short (compared to $\lambda/2$) channels that emerge from the ends of the trigger dipole or even a plasma loop that connects the ends of the trigger during the experimental time τ. We shall consider a self-sustained dis-

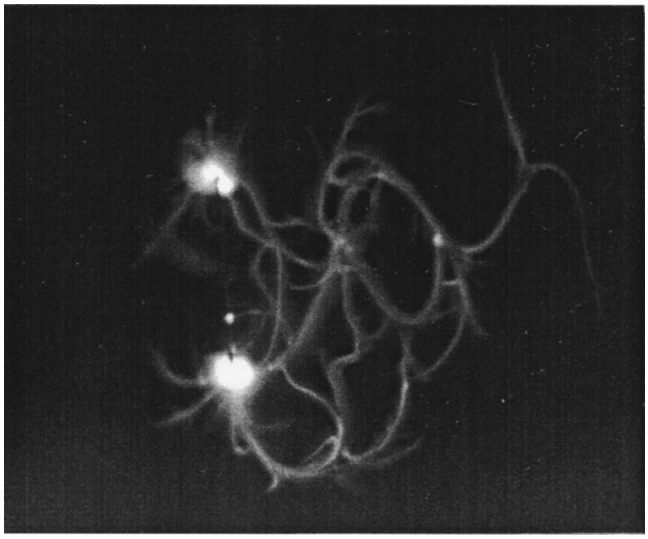

FIG. 3. Self-sustained developed triggered microwave streamer discharge with low subcriticality: $\tau=40$ μs, $2L=3$ cm, $2a=8\times10^{-2}$ cm.

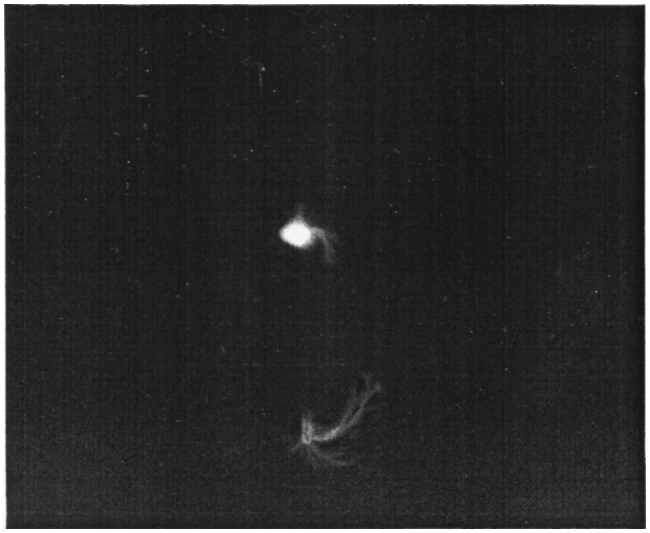

FIG. 4. A highly subcritical microwave discharge in air, which is attached to the trigger: $\tau=40$ μs, $2L=3$ cm, $2a=8\times10^{-2}$ cm.

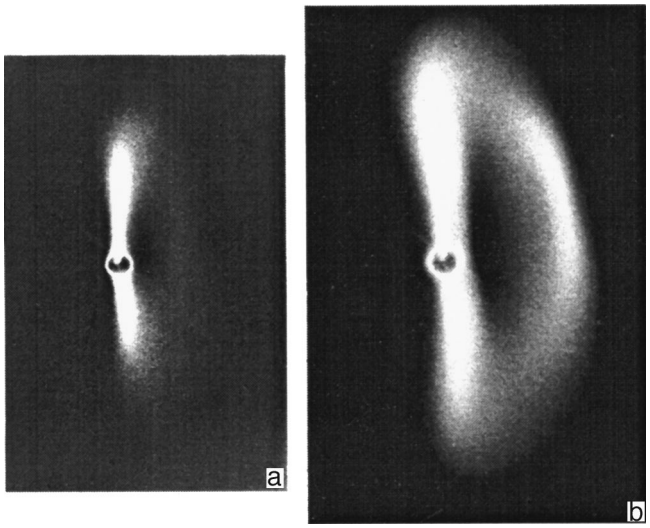

FIG. 5. Diffuse microwave discharge initiated by a small sphere in low pressure air.

charge to be one which, over the 40 μs experimental time, could form plasma resonance dipoles with lengths close to $\lambda/2$ that were not in direct contact with the trigger.[3]

A typical example of a subcritical discharge capable of independent development for a field amplitude E_0 close to the limiting value separating microwave discharges capable of independent development from those that are attached to the trigger is shown in Fig. 7 ($2L=2$ cm, $p=150$ Torr, and $E_0=1.2$ kV/cm). A ''plasma sinusoid'' extending from the lower end of the trigger toward the microwave source can be seen in this figure. This figure illustrates clearly the self development principle for a microwave streamer discharge in a subcritical field. Each of the comparatively straight segments of this sinusoid near the trigger, with an inclination of roughly 45° to $\mathbf{E}_0$ and a length of roughly $\lambda/2$, is essentially a plasma trigger which ensures the subsequent development of the discharge.[1]

FIG. 6. Subcritical microwave discharge with streamer loops ''tied'' to the initiating trigger: $\tau=40$ μs, $2a=8\times10^{-2}$ cm.

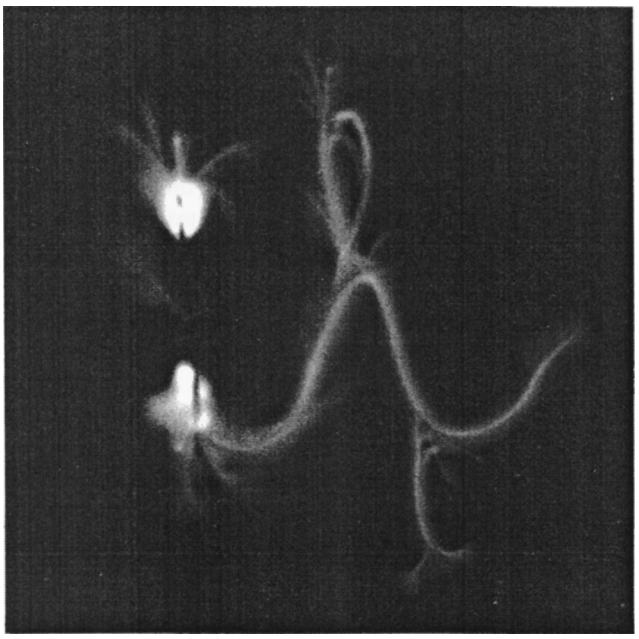

FIG. 7. Self-sustained, developed subcritical triggered microwave discharge at a low field level: $\tau=40$ μs, $2a=8\times10^{-2}$ cm.

The length of the streamer sinusoid measured from Fig. 7 was about 10 cm. Thus, the average rate of growth of the streamer, V_{str}, is at least 2.5×10^5 cm/s. This is essentially the same as V_{str} estimated above from Fig. 3 for a discharge with a developed spatial structure for low subcriticality.

An empirical formula for V_{str} was given in Ref. 4. It implies that as E_0 is reduced, the average velocity of a microwave streamer approaches a constant value V_{min}, which depends only on the air pressure. For example, for $p=150$ Torr this formula gives $V_{min}=3.7\times10^5$ cm/s, which, given the spatial uncertainty in the trajectory of the streamer in Fig. 7 and the spatial nonuniformity of E_0 in the focal region of the electromagnetic beam, is in fair agreement with experiment. That is, if the microwave streamer broke away from the trigger and began to develop independently, then it might have a velocity V_{str} at least equal to V_{min}. In this sense, an attached streamer discharge is obviously a fundamentally different plasma formation and there should be a distinct boundary with respect to E_0 between the two forms of subcritical discharge. In the following we shall denote this limiting field by E_L. The experimentally determined E_L will also be the lower limit for using the formula for V_{fr} (or V_{str}) given in Ref. 4.

Figure 8 is a plot of the minimum E_0 for each experimental value of p such that a self sustained microwave discharge exists: ● corresponds to a trigger dipole with $2L=3$ cm, and × to one with $2L=2$ cm. These values were obtained from an analysis of Figs. 1 and 2 employing the criterion formulated above for external differences between the two forms of discharge.

It is clear from Fig. 8 that the experimental values of E_L for the longer trigger dipole are somewhat lower. At the same time, in terms of the meaning of a limiting field, it is a characteristic of the discharge itself and should not depend on the method of initiating it. Thus, the weak dependence of

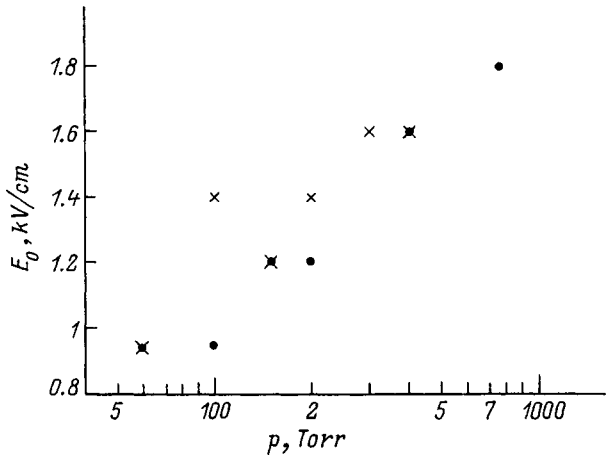

FIG. 8. Minimum limiting field amplitude for which a self-sustained, developed subcritical microwave discharge exists ($2a=8\times10^{-2}$ cm).

the field E_L on $2L$ found here is indicative of the degree of uncertainty in its measurement in this particular experiment.

In Fig. 8, p is on a logarithmic scale, while E_0 is on a linear scale. In this plot, the experimental points fit a straight line fairly well.

The function $\Psi(p)$ approximating the data of Fig. 8 is plotted in Fig. 9. The region above this curve is the region of attached subcritical microwave discharges and that below the curve, of self sustained discharges. The region below the line $\Psi=1$ corresponds to supercritical discharges. It is evident from this figure that as p increases, there is an increase in the range of Ψ within which a self sustained subcritical microwave discharge can exist. Under the experimental conditions it is a large as $1\leqslant\Psi\leqslant17$.

Extrapolating the $\Psi(p)$ curve to lower p gives an intersection with the level line $\Psi=1$ at $p=20-30$ Torr. This is in good agreement with the experimental data of Ref. 2. There it was stated that in an electromagnetic beam with $\lambda=8.5$ cm, a streamer discharge in air can only develop for $p\geqslant25$ Torr. At lower p, a diffuse microwave discharge essentially does not extend outside the region of the electromagnetic beam with $E_0\geqslant E_k$, i.e., it cannot develop in a subcritical field.

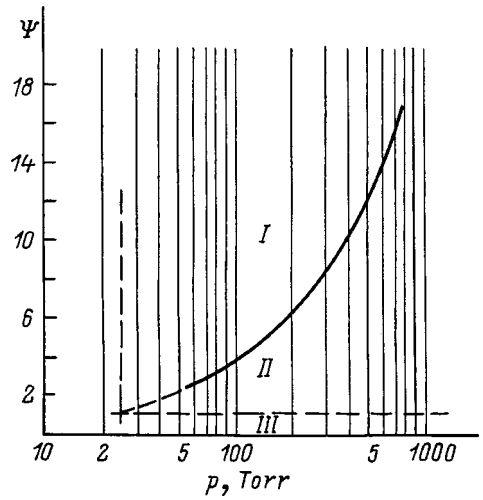

FIG. 9. Regions in which the different types of microwave streamer discharge exist (*I–III*).

At present there is no quantitative explanation for the dependence $\Psi(p)$ shown in Fig. 9. At the same time, it is clear that the way to construct a suitable model involves an analysis of the plasma and electromagnetic processes taking place in an individual microwave streamer trigger dipole. As Fig. 7 shows, it is the major, simplest structure-forming element of the discharge. Evidently, the dissipative processes in a trigger dipole of this sort determine the amplitude of the microwave current along it and, therefore, the charge at its growing end. The curvature of the plasma boundary at this end, on the other hand, determines the surface charge density there and, therefore, the field amplification at the head of the streamer. The increase in the size of the enhanced field region at the end of the plasma trigger dipole as its length approaches the resonance value, in turn, creates the conditions for generation of a new microwave streamer trigger dipole.

The experiments have shown, therefore, that the pulsed subcritical streamer discharge in air initiated by a microwave dipole in a electromagnetic *TEM*-beam in the centimeter wavelength range develops in two fundamentally different forms, depending on the degree of subcriticality of the initial field. For low subcriticality, a discharge generated at the ends of the trigger dipole separates from it and develops self consistently, propagating counter to the microwave radiation in the form of lengthening, branching plasma streamer channels. The average rate at which they grow cannot be below a few times 10^5 cm/s. For large subcriticalities, the discharge cannot break away from the trigger dipole and remains as if attached to the trigger dipole throughout a microwave pulse lasting tens of microseconds. A lower bound on the field for existence of a self-sustained (independently developing) discharge has been determined in these experiments. This result implies that both the level of the limiting field and the range of subcriticality within which such a discharge can exist increase as the air pressure is raised. Under the experimental conditions at atmospheric pressure, the maximum subcriticality was 17. Therefore, in this case, with other conditions the same, the power of the generator required to create a microwave streamer discharge is almost 300 times lower than that for an analogous supercritical discharge. These experiments have confirmed that the existence of a self-sustained microwave discharge is related specifically to its streamer structure, which is only characteristic of comparatively high air pressures. Under the experimental conditions, they can be produced only at pressures above 20 Torr. This pressure is also a limit below which a self-sustained subcritical discharge cannot develop. Finally, the data presented here can be used to establish more precisely the role of individual physical factors when constructing a model of microwave streamer discharges.

We thank K. V. Aleksandrov for help in doing the experiments.

[1] L. P. Grachev, I. I. Esakov, and K. V. Khodataev, Zh. Tekh. Fiz. **68**(4), 33 (1968) [Tech. Phys. **43**, 378 (1998)].
[2] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **64**(1), 74 (1994) [Tech. Phys. **39**, 40 (1994)].
[3] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Pis'ma Zh. Tekh. Fiz. **18**(22), 34 (1992) [Sov. Tech. Phys. Lett. **18**(11), 737 (1992)].
[4] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **65**(5), 21 (1995) [Tech. Phys. **40**, 416 (1995)].

Translated by D. H. McNeill

Criterion for microwave breakdown of gases

V. A. Lisovskiĭ

Kharkov State University, 310077 Kharkov, Ukraine
(Submitted May 27, 1998; resubmitted September 8, 1998)
Zh. Tekh. Fiz. **69**, 25–29 (November 1999)

An analytical criterion is obtained for microwave breakdown of gases that takes into account the different diffusion coefficients of electrons along and perpendicular to the direction of the electric field. Discharge ignition without electron attachment to the gas molecules and in electronegative gases is examined, along with microwave discharge ignition when a weak dc electric field is applied. © *1999 American Institute of Physics.* [S1063-7842(99)00511-5]

INTRODUCTION

Gaseous discharges in microwave electric fields are widely used in various areas of technology. Because of the rapid development of microwave electronics, considerable attention is devoted to studying the conditions for discharge production in microwave transmission lines, since these conditions set the power limits for these lines. Recently, microwave discharges have also found applications in various engineering processes: etching of semiconductor materials, oxidation of silicon and aluminum surfaces, deposition of thin films, pumping gas lasers, etc. Interest in microwave gas discharges also arises in connection with the possibility of obtaining new data on the interactions of electrons with atoms and ions. Thus, it is not surprising that many experimental and theoretical papers deal with the ignition of microwave discharges.[1–15]

It is known that, in an alternating electric field, charged particles undergo oscillatory motion whose amplitude decreases as the frequency of the field rises. At microwave frequencies the amplitude of the electron oscillations is much smaller than the size of the discharge volume (cavity, waveguide, etc.). Here the bulk of the electron cloud does not reach the walls of the discharge volume during the oscillations in the microwave field. The heavy positive (and negative) ions are even less able to reach the boundaries of the discharge volume, since the amplitude of the displacements of the ions in a microwave field is thousands of times smaller than the corresponding amplitude for the electrons. Thus, in a microwave discharge the only charge carriers are the electrons, while the ions are practically uninvolved in igniting a discharge. The small amplitude of the electron displacements in a microwave field means, first of all, that wall processes essentially play no role and, second, that the displacement of the electrons in the field cannot produce space charge effects within a significant volume (as opposed to discharges in dc or rf electric fields).[16]

The condition for development of a self-sustained discharge is equality of the number of electrons created in the discharge volume to the number leaving it without an external ionization source,

$$\frac{\partial n_e}{\partial t}=0, \tag{1}$$

where n_e is the electron density and t is time.

If a microwave discharge is ignited when processes at the discharge vessel wall can play no significant role, then the only process for electron production is electron-impact ionization of gas atoms. Free electrons can be lost from the discharge volume by diffusing to the discharge vessel walls and by attachment to molecules of an electronegative gas, and when microwave and dc electric fields are present simultaneously, electrons can also leave the discharge volume through directed motion toward the cavity walls.

In the simplest case, when electron-impact ionization of the gas molecules and diffusion of electrons to the cavity wall predominate, the criterion for microwave breakdown of the gas has the form[1–6]

$$\frac{\nu_i}{D_e}=\frac{1}{\Lambda^2}. \tag{2}$$

Here ν_i is the electron-impact ionization rate for the gas molecules, D_e is the electron diffusion coefficient, and Λ is the diffusion length, which depends on the geometry of the discharge vessel. For example, for a cylindrical cavity with radius R and height L, the diffusion length is given by

$$\frac{1}{\Lambda^2}=\left(\frac{2.4}{R}\right)^2+\left(\frac{\pi}{L}\right)^2. \tag{3}$$

Note that Eqs. (2) and (3) are derived under the assumption that electron diffusion is isotropic, i.e., the electron diffusion coefficients along and perpendicular to the electric field are the same. However, later experimental and theoretical work[17–22] has shown that, in general, the electron diffusion coefficient along the direction of the electric field, D_L, is not equal to that perpendicular to the field, D_e; that is, the electron diffusion is anisotropic.

In this paper a criterion for microwave breakdown in gases is obtained which accounts for the anisotropy in the electron diffusion. The ignition of microwave discharges in electropositive and electronegative gases is considered, as well the ignition of combination (microwave + weak dc electric fields) discharges.

MICROWAVE BREAKDOWN WITHOUT ELECTRON ATTACHMENT

In this case we shall assume that the only electron loss process is diffusion to the discharge vessel walls. Suppose the discharge volume is filled with a gas whose molecules cannot capture free electrons.

We consider a gas filled gap in a cylindrical cavity of height L and radius R. The z axis is directed along the cavity axis and the radial coordinate is perpendicular to the axis, while the origin is located at the center of the discharge gap. We shall consider the ignition of the discharge in the case where the cavity is excited by a TM_{010} electromagnetic wave. For a wave of this type, the resonant frequency is independent of the cavity length and the electric field is parallel to its axis.[11] We write down the balance equation for the electrons in the form

$$\frac{\partial n_e}{\partial t}=\nu_i n_e+D_e\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial n_e}{\partial r}\right)+D_L\frac{\partial^2 n_e}{\partial z^2}. \tag{4}$$

We shall use the following boundary conditions:

$$n_e(R,z)=0, \tag{5}$$

and

$$n_e\left(r,\pm\frac{L}{2}\right)=0. \tag{6}$$

The first term on the right hand side of Eq. (4) is the rate of charged particle production owing to electron-impact ionization of the gas molecules, the second term is the rate of loss of electrons perpendicular to the electric field at the radial walls of the cylinder, and the third is the rate of loss of electrons along the direction of the electric field. Since the necessary condition for ignition of the discharge is that the rates of electron production and loss be equal, Eq. (4) takes the form

$$\frac{\nu_i}{D_e}+\frac{1}{n_e}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial n_e}{\partial r}\right)+\frac{D_L}{D_e}\frac{1}{n_e}\frac{\partial^2 n_e}{\partial z^2}=0. \tag{7}$$

We solve Eq. (7) by separating the variables r and z. The electron density is written as the product of two functions, one of which depends only on the radial coordinate r, while the other depends only on z, i.e.,

$$n_e(r,z)=f(r)\cdot\varphi(z). \tag{8}$$

Then we obtain the following expression for the electron density:

$$n_e(r,z)=A_0J_0\left(\frac{2.4}{R}r\right)\cos\left(\frac{\pi}{L}z\right), \tag{9}$$

where A_0 is a constant and J_0 is the Bessel function of zero order.

The criterion for microwave breakdown of the gas, including the anisotropy of the electron diffusion, takes the form

$$\frac{\nu_i}{D_e}=\left(\frac{2.4}{R}\right)^2+\frac{D_L}{D_e}\frac{\pi^2}{L^2}. \tag{10}$$

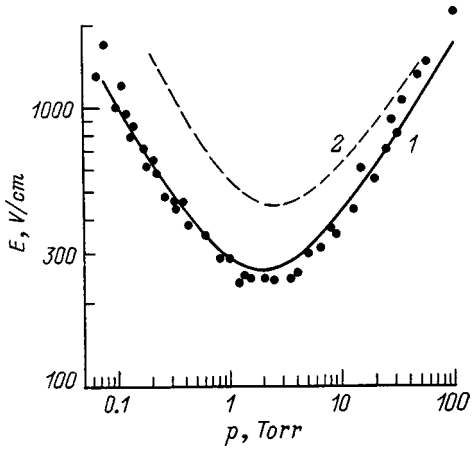

FIG. 1. Microwave breakdown curves for xenon: $f=2800$ MHz, $L=0.318$ cm, $R=3.67$ cm; ● — experiment,[9] *1* — calculation according to Eq. (10), *2* — calculation according to Eqs. (2) and (3).

Therefore, taking the anisotropy of electron diffusion into account makes the diffusion length Λ a function of the geometric dimensions of the cavity, but it also depends on the ratio of the diffusion coefficients along and perpendicular to the direction of the electric field. In general, $D_e\neq D_L$ and D_L/D_e is a complicated nonmonotonic function of the ratio E/p of the electric field and gas pressure. Thus, it is very difficult to solve Eq. (10) analytically and in order to find the breakdown field it has to be solved numerically. Figures 1 and 2 show experimental values of the microwave breakdown field for xenon[9] and hydrogen,[2] as well as the breakdown curves calculated according to Eqs. (2), (3), and (10). Here we have used the electron transport coefficients given in Refs. 23–26. It is evident from these figures that our theoretical curves calculated using Eq. (10) are in satisfactory agreement with the experimental data. The theoretical values of the microwave breakdown field given by Eqs. (2) and (3) for isotropic diffusion are too high compared to the experimental data and the calculations according to Eq. (10). This is because, in most cases, $D_L<D_e$, i.e., electron loss by diffusion along the electric field direction is much less than

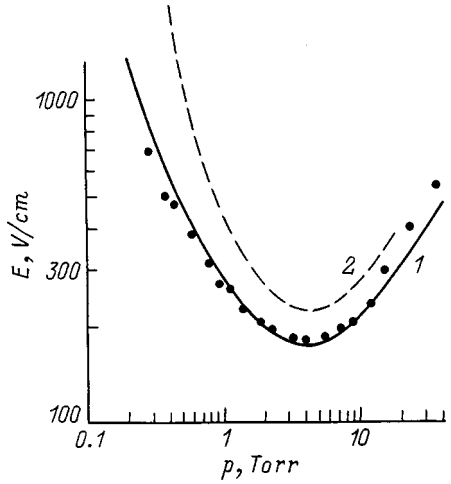

FIG. 2. Microwave breakdown curves for hydrogen: $f=3000$ MHz, $L=2.54$ cm, $R=4.07$ cm; ● — experiment,[2] *1, 2* — as in Fig. 1.

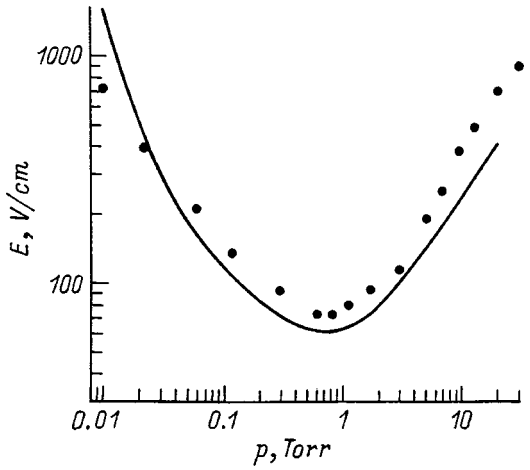

FIG. 3. Microwave breakdown curves for oxygen: $f=994$ MHz, $L=4.74$ cm, $R=4$ cm; ● — experiment,[10] the curve is calculated using Eq. (12).

across the field (to the radial walls of the cavity). Since the diffusion coefficients are assumed equal in Eqs. (2) and (3), i.e., $D_L=D_e$, here the electron losses appear to be greater than they actually are. Thus, the theoretical curve based on Eqs. (2) and (3) lies significantly above the experimental points.

In some cases, when it can be assumed that $D_e \approx D_L$ (e.g., for CO_2[17,22]), the modified microwave breakdown criterion (10) can be reduced to the conventional criterion of Eqs. (2) and (3).

MICROWAVE BREAKDOWN IN ELECTRONEGATIVE GASES

When a microwave discharge is ignited in an electronegative gas, the electron density obeys the equation

$$\frac{\partial n_e}{\partial t}=(\nu_i-\nu_a)n_e+D_e\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial n_e}{\partial r}\right)+D_L\frac{\partial^2 n_e}{\partial z^2}, \qquad (11)$$

where ν_a is the rate at which free electrons are attached to the electronegative gas molecules.

Then it is easy to show that, in this case, the criterion for microwave breakdown of the gas takes the form

$$\frac{\nu_i}{D_e}=\frac{\nu_a}{D_e}+\left(\frac{2.4}{R}\right)^2+\frac{D_L}{D_e}\frac{\pi^2}{L^2}. \qquad (12)$$

For large discharge vessels, the loss of electrons to the walls by diffusion is low, and the dominant electron loss process is electron attachment. In this case, the microwave breakdown criterion (12) simplifies to[16]

$$\nu_i=\nu_a. \qquad (13)$$

Therefore, a self-sustained microwave discharge can be ignited in a large cavity when the rate of electron-impact ionization of the gas molecules becomes equal to the rate of electron attachment to the gas molecules.

Figure 3 shows a theoretical plot of the microwave breakdown curve for oxygen obtained using the criterion (12), together with experimental data.[10] In the calculations we have used the transport coefficients for electrons in oxygen given in Refs. 26 and 27 and the electron attachment coefficients for oxygen molecules given in Ref. 27. Note that there is a large scatter in the reported measurements of the electron attachment coefficient.[27] This uncertainty in the electron attachment coefficient for oxygen molecules could be the cause of the discrepancy between our theoretical curve and the experimental breakdown curve.[10]

EFFECT OF A WEAK DC ELECTRIC FIELD ON MICROWAVE BREAKDOWN OF GASES

In this section we derive the criterion for breakdown of a gas in a combination (microwave + weak dc electric field) discharge. Let a microwave field E and a weak (too small to contribute to the ionization of the gas) dc electric field E_{dc} be created simultaneously in a cavity with the directions of the two fields coincident (a longitudinal combination discharge). The balance equation for the electrons is

$$\frac{\partial n_e}{\partial t}=\nu_i n_e+D_e\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial n_e}{\partial r}\right)+D_L\frac{\partial^2 n_e}{\partial z^2}-V_{dc}\frac{\partial n_e}{\partial z}, \qquad (14)$$

where $V_{dc}=eE_{dc}/m\nu_{en}$ is the drift velocity of the electrons in the dc electric field, ν_{en} is the electron-neutral collision frequency, and e and m are the charge and mass of the electrons, respectively.

The boundary conditions are taken to be Eqs. (5) and (6). We seek a solution to Eq. (14) by separation of variables. In the steady-state case, the electron density is given by

$$n_e(r,z)=A_1 J_0\left(\frac{2.4}{R}r\right)\exp\left(\frac{V_{dc}}{2D_L}z\right)\cos\left(\frac{\pi}{L}z\right), \qquad (15)$$

where A_1 is a constant.

The criterion for breakdown of a gas in microwave and weak dc electric fields (longitudinal combination electric field) is

$$\frac{\nu_i}{D_e}=\left(\frac{2.4}{R}\right)^2+\frac{D_L}{D_e}\frac{\pi^2}{L^2}+\frac{V_{dc}^2}{4D_eD_L}. \qquad (16)$$

In the absence of a dc electric field ($E_{dc}=0$), Eq. (16) yields the criterion (10) for microwave breakdown of the gas. When the electron diffusion can be assumed to be isotropic ($D_e \approx D_L$), the criterion (16) reduces to[4,28,29]

$$\frac{\nu_i}{D_e}=\left(\frac{2.4}{R}\right)^2+\frac{\pi^2}{L^2}+\left(\frac{V_{dc}}{2D_e}\right)^2. \qquad (17)$$

Figure 4 shows microwave breakdown curves with a weak applied dc electric field E_{dc}, calculated using Eq. (16). Evidently, as E_{dc} is increased, there is a rise in the microwave breakdown field, in accord with Ref. 29. Note that Eqs. (16) and (17) are valid only for a weak dc electric field which does not contribute to ionization of the gas and generate positive ion fluxes toward the cathode (and, as a consequence, cause significant ion and electron emission from the cathode surface). Of course, in stronger electric fields E_{dc}, which do contribute to ionization of the gas, the breakdown microwave field begins to decrease and approaches zero when E_{dc} equals the ignition potential

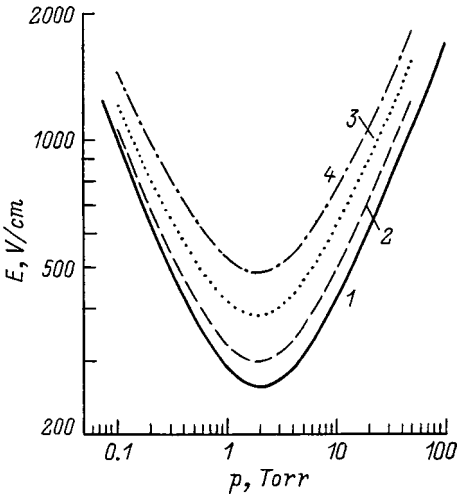

FIG. 4. Microwave breakdown curves for xenon with $E_{dc}=0$ (*1*), 50 (*2*), 100 (*3*), and 150 V/cm (*4*). Calculated according to Eq. (16); *f*, *L*, and *R* are as in Fig. 1.

for a dc glow discharge. Unfortunately, there are absolutely no published experimental ignition curves for microwave discharges with an applied electric field. Thus, we cannot compare our calculated curves with experimental data.

CONCLUSION

A criterion for the microwave breakdown of a gas has been derived here that includes the difference in the electron diffusion coefficients along and perpendicular to the direction of the electric field. The ignition of microwave discharges without electron attachment to the gas molecules and in electronegative gases has been examined, as well as the ignition of longitudinal combination discharges (microwaves + a weak dc electric field).

[1] M. A. Herlin M.A. and S. C. Brown, Phys. Rev. **74**, 291 (1948); *ibid*, p. 1650.
[2] A. D. MacDonald and S. C. Brown, Phys. Rev. **75**, 411 (1949); **76**, 1634 (1949).
[3] S. C. Brown and A. D. MacDonald, Phys. Rev. **76**, 1629 (1949).
[4] S. C. Brown, Proc. IRE **39**, 1493 (1951).
[5] T. Kihara, Rev. Mod. Phys. **24**, 45 (1952).
[6] W. P. Allis and S. C. Brown, Phys. Rev. **87**, 419 (1952).
[7] H. H. Oskam, J. Appl. Phys. **27**, 848 (1956).
[8] D. J. S. Rose and S. C. Brown, J. Appl. Phys. **28**, 561 (1957).
[9] H. M. Bradford, D. M. Fraser, G. F. O. Langstroth, and A. D. MacDonald, Can. J. Phys. **37**, 1166 (1959).
[10] A. D. MacDonald, D. U. Gaskell, and N. H. Gitterman, Phys. Rev. **130**, 1841 (1963).
[11] A. D. MacDonald, *Microwave Breakdown in Gases* [Wiley, New York (1966); Mir, Moscow (1969), 212 pp.].
[12] A. G. Pokreev and A. G. Reuka, Ukr. Fiz. Zh. **29**, 1788 (1984).
[13] G. A. Anashkin, Radiotekh. Élektron., No. 12, pp. 2576–2584 (1987).
[14] A. L. Vikharev, O. A. Ivanov, O. Yu. Kuznetsov, and A. N. Stepanov, Fiz. Plazmy **13**, 1124 (1987) [Sov. J. Plasma Phys. **13**, 648 (1987)].
[15] Yu. P. Raizer, *Physics of Gaseous Discharges* [in Russian], Nauka, Moscow (1987), 592 pp.
[16] V. E. Golant, Usp. Fiz. Nauk **65**, 39 (1958).
[17] E. B. Wagner, F. J. Davis, and G. S. Hurst, J. Chem. Phys. **47**, 3138 (1967).
[18] J. H. Parker and J. J. Lowke, Phys. Rev. **181**, 290 (1969).
[19] J. J. Lowke and J. H. Parker, Phys. Rev. **181**, 302 (1969).
[20] H. R. Skullerud, J. Phys. B **2** 696 (1969).
[21] R. E. Robson, Aust. J. Phys. **25**, 685 (1972).
[22] L. G. H. Huxley and R. W. Crompton, *The Diffusion and Drift of Electrons in Gases* [Wiley, New York (1974); Mir, Moscow (1977), 672 pp.].
[23] M. Suzuki, T. Taniguchi, N. Yoshimura, and H. Tagashira, J. Phys. D **25**, 50 (1992).
[24] H. A. Blevin, J. Fletcher, and S. R. Hunter, J. Phys. D **9**, 1671 (1976); **11**, 1653 (1978).
[25] H. T. Saelee and J. Lucas, J. Phys. D **10**, 343 (1977).
[26] W. Roznerski and K. Leja, J. Phys. D **13**, L181 (1980); **17**, 279 (1984).
[27] M. S. Naidu and A. N. Prasad, J. Phys. D **3**, 954 (1970).
[28] L. J. Varnerin and S. C. Brown, Phys. Rev. **79**, 946 (1950).
[29] S. C. Brown, in *Handbuch der Physik*, Vol. 22, Springer-Verlag, Berlin (1956), pp. 531–575.

Translated by D. H. McNeill

Potential surface waves at the boundary of a metal with a magnetoactive plasma at finite pressure

N. A. Azarenkov, A. N. Kondratenko, and Yu. O. Tyshetskiĭ

Kharkov State University, 310077 Karkov, Ukraine
(Submitted November 24, 1997)
Zh. Tekh. Fiz. **69**, 30–33 (November 1999)

The propagation of surface-type potential waves along the interfacial boundary of a plasma with an ideally conducting metal in an external magnetic field perpendicular to the boundary is examined. It is shown that a necessary condition for the existence of these waves in the system is a finite gas kinetic pressure. Dispersion relations for these waves and expressions for the penetration depth of the wave fields into the plasma are obtained, and they are studied numerically for various plasma parameters. The frequency region for propagation of these waves is found. It is also shown that in a nonzero external magnetic field a system of this kind has a range of frequencies in which the wave is a generalized surface wave. © *1999 American Institute of Physics.* [S1063-7842(99)00611-X]

The properties of surface waves in plasma–metal structures have recently been under intense study.[1,2] Interest in these structures arises in the course of research on plasma and semiconductor electronics,[3,4] limiters and divertors in toroidal controlled fusion devices,[5] the construction of plasma sources,[6–8] etc. The dispersion properties and spatial distributions of surface wave fields in plasma–metal structures have been studied in some detail, both theoretically and experimentally.[1] Here we have in mind surface waves produced by collective excitations in a plasma medium, whose frequency is substantially below that of the inherent electronic perturbations in a metal. Results from a study of surface waves at the boundary of a metal with a free and magnetoactive plasma have been presented previously.[1] The properties of potential and nonpotential surface waves in a magnetoactive plasma were studied in the case where the external magnetic field is parallel to the interface. In this paper we show that surface-type waves can also propagate along the boundary of a plasma with an ideally conducting metal when an external magnetic field is perpendicular to the interface. This geometry of the problem is typical of processes for plasma working of conducting surfaces, when it is necessary to establish the transverse dimensions of a flow, or in divertor plasmas. Here a necessary condition for the existence of a surface wave in a structure of this sort is a finite gas kinetic pressure. In a cold plasma, waves of this type cannot exist.

We choose a coordinate system such that the plasma occupies the region $x>0$ and has a boundary with an ideally conducting surface ($\sigma\to\infty$) in the plane $x=0$. The interface between the metal and plasma is assumed to be sharp. The model of a sharp boundary is valid if the penetration depth of the surface wave field is much greater than the dimensions of the transition region. An external magnetic field is directed along the x axis. All of the wave perturbations depend on position and time as

$$\mathbf{A}(\mathbf{R},t)=A(x)\exp[i(k_2 y-\omega t)]. \tag{1}$$

The system of equations for this wave process consists of the Poisson equation for the potential φ of the field and a system of linearized quasihydrodynamic equations with cold ions and hot electrons:[2]

$$\nabla^2\varphi=4\pi e(n_e-n_i),$$

$$\frac{\partial \mathbf{V}_\alpha}{\partial t}=\frac{e_\alpha}{m_\alpha}\nabla\varphi+\frac{e_\alpha}{m_\alpha c}(\mathbf{V}_\alpha\times\mathbf{H}_0)-v_{T\alpha}^2\frac{\nabla n_\alpha}{n_{0\alpha}},$$

$$\frac{\partial n_\alpha}{\partial t}+n_{0\alpha}(\nabla\cdot\mathbf{v}_\alpha)=0. \tag{2}$$

Here e_α, m_α, $n_{0\alpha}$, $v_{T\alpha}$, n_α, $\mathbf{v}_\alpha$ are the charge, mass, equilibrium density, thermal velocity, perturbation density, and hydrodynamic velocity of the particles of species α ($\alpha=e,i$, respectively for electrons and ions), and φ is the perturbation potential. Since the ions are assumed cold in our model, we have $v_{Ti}=0$, while $v_{Te}=\sqrt{T_e/m_e}$ (T_e is the electron temperature in energy units). The plasma is assumed homogeneous with respect to the density and quasineutral ($n_{0e}=n_{0i}\equiv n_0$).

If we seek a solution for the system of Eqs. (2) in the form of waves traveling along the y axis (1), then it is possible to obtain the following differential equation for the potential distribution in the plasma:

$$\frac{d^4\varphi}{dx^4}+\left(\frac{\omega^2}{v_{Te}^2}\frac{\varepsilon_1}{\varepsilon_{i1}}-k_y^2\left(\frac{\varepsilon_{i2}}{\varepsilon_{i1}}-\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)\right)\frac{d^2\varphi}{dx^2}+k_y^2\left(-\frac{\omega^2}{v_{Te}^2}\frac{\varepsilon_2}{\varepsilon_1}+\frac{\omega^2}{\omega^2-\omega_{ce}^2}\frac{\varepsilon_{i2}}{\varepsilon_{i1}}k_y^2\right)\varphi=0, \tag{3}$$

where

$$\varepsilon_1=\varepsilon_{i1}-\frac{\omega_{pe}^2}{\omega^2},\quad \varepsilon_{i1}=1-\frac{\omega_{pi}^2}{\omega^2},$$

$$\varepsilon_2=\varepsilon_{i2}-\frac{\omega_{pe}^2}{\omega^2-\omega_{ce}^2},\quad \varepsilon_{i2}=1-\frac{\omega_{pi}^2}{\omega^2-\omega_{ci}^2},$$

$$\omega_{p\alpha}=\left(\frac{4\pi e_\alpha^2 n_0}{m_\alpha}\right)^{1/2},\quad \omega_\alpha=\frac{e_\alpha H_0}{m_\alpha c}.$$

If we set $H_0=0$ in Eq. (3), then we obtain an equation for the potential distribution in a free, hot plasma.[2]

The solution of Eq. (3) for the potential has the form

$$\varphi(x)=A_1\exp(-\lambda_1 x)+A_2\exp(-\lambda_2 x),\quad \lambda_{1,2}>0, \tag{4}$$

where

$$\lambda_{1,2}^2=\frac{1}{2}\left\{k_y^2\left(\frac{\varepsilon_{i2}}{\varepsilon_{i1}}+\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)-\frac{\omega^2}{v_{Te}^2}\frac{\varepsilon_1}{\varepsilon_{i1}}\pm\sqrt{\left(\frac{\omega^2}{v_{Te}^2}\frac{\varepsilon_1}{\varepsilon_{i1}}-k_y^2\left(\frac{\varepsilon_{i2}}{\varepsilon_{i1}}+\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)\right)^2-4k_y^2\left(\frac{\omega^2k_y^2}{\omega^2-\omega_{ce}^2}\frac{\varepsilon_{i2}}{\varepsilon_{i1}}-\frac{\omega^2}{v_{Te}^2}\frac{\varepsilon_2}{\varepsilon_{i1}}\right)}\right\}. \tag{5}$$

In the following we shall study wave perturbations of a surface type caused by electron perturbations, i.e., with $\omega \gg \omega_{pi}, \omega_{ci}$. Then $\varepsilon_{i1}=\varepsilon_{i2}=1$, $\varepsilon_1=1-\omega_{pe}^2/\omega^2$, and $\varepsilon_2=1-\omega_{pe}^2/(\omega^2-\omega_{ce}^2)$ and the expressions for $\lambda_{1,2}^2$ can be written in the form

$$\lambda_{1,2}^2=\frac{1}{2}\frac{\omega^2}{v_{Te}^2}\left\{\frac{v_{Te}^2k_y^2}{\omega^2}\left(1+\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)-\varepsilon_1\pm\sqrt{\varepsilon_1^2+\frac{v_{Te}^2k_y^2}{\omega^2}\left(1-\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)\left[\frac{v_{Te}^2k_y^2}{\omega^2}\left(1-\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)+2\left(1+\frac{\omega_{pe}^2}{\omega^2}\right)\right]}\right\}. \tag{6}$$

If we set $H_0=0$ in Eq. (6), then $\lambda_1^2=k_y^2$, while $\lambda_2^2=k_y^2-\omega^2\varepsilon_1/v_{Te}^2$, i.e., the expressions for the reciprocals of the penetration depths of the surface waves into the plasma will be the same as the corresponding expressions given in Ref. 9 for potential surface wave in a free plasma.

In order to obtain a dispersion relation for the potential surface waves being studied here, the initial system of Eqs. (2) must be supplemented by boundary conditions. The system of boundary conditions for a finite pressure plasma is[2]

$$\varphi(x=0)=0,\quad v_{ex}(x=0)=0. \tag{7}$$

The system of Eqs. (2) implies that the boundary condition for the normal component of the hydrodynamic velocity of the electrons can be written in the form

$$\frac{e}{m}\frac{d\varphi}{dx}-\frac{v_{Te}^2}{4\pi e n_0}\frac{d}{dx}\left(\frac{d^2\varphi}{dx^2}-k_y^2\varphi\right)\Bigg|_{x=0}=0. \tag{8}$$

On substituting Eq. (4) for the potential and its derivatives in the system of boundary conditions (7) and (8), the compatibility conditions for the system yield a dispersion relation for the potential surface waves,

$$1+\frac{v_{Te}^2}{\omega_{pe}^2}k_y^2=\frac{v_{Te}^2}{\omega_{pe}^2}(\lambda_1^2+\lambda_1\lambda_2+\lambda_2^2). \tag{9}$$

Its solution has the form

$$k_y^2=\frac{1}{v_{Te}^2}\left(\frac{\omega^2}{\omega_{ce}^2}-1\right)\{\omega^2+\omega_{ce}^2+\omega_{pe}^2-\sqrt{(\omega^2-\omega_{ce}^2+\omega_{pe}^2)^2+4\omega_{ce}^2\omega_{pe}^2}\}. \tag{10}$$

The wave propagation condition $k_y^2>0$ implies that $\omega>\omega_{ce}$ is the region for existence of surface waves propagating perpendicular to the external magnetic field. A necessary condition for the existence of this surface wave is a finite electron velocity. In the limit of $H_0=0$, the expression for k_y^2 transforms to the known solution for potential surface waves in a free plasma[1,9]

$$\omega\approx\omega_{pe}\sqrt{|k_y|r_{De}},\quad r_{De}=\frac{v_{Te}}{\omega_{pe}}. \tag{11}$$

With a small, but finite magnetic field, when the conditions $\omega_{ce}^2\ll\omega^2\ll\omega_{pe}^2$ are satisfied, the expression for the eigenfrequency of the surface wave transforms to

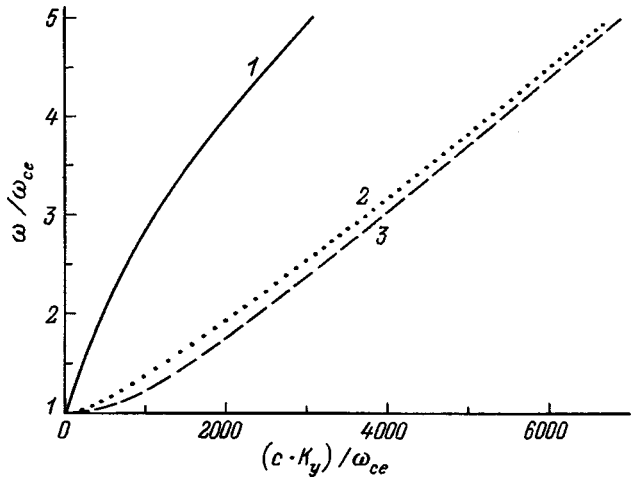

FIG. 1. Dispersion for different values of $\Omega\equiv\omega_{pe}/\omega_{ce}$ (the external magnetic field is fixed). The plasma temperature $T_e=1000$ K; $\Omega=10$ (*1*), 1 (*2*), 0.1 (*3*).

$$\omega \approx \omega_{pe}\sqrt{|k_y| r_{De}}\left(1+\frac{\omega_{ce}^2}{4\omega_{pe}^2}\right). \tag{12}$$

It can be seen from this expression that the eigenfrequency of the surface wave increases as the external magnetic field is raised.

The results of a numerical analysis of the dispersion relation (10) for different values of the parameter $\Omega = \omega_{pe}/\omega_{ce}$ (for a fixed external magnetic field H_0) are shown in Fig. 1. The curves were obtained for the following plasma parameters: $T_e = 1000$ K and $H_0 = 100$ G. It is evident from this figure that for low magnetic field strengths H_0 the dispersion curve transforms to the curve of Ref. 1 for a free plasma. As Ω increases, the phase and group velocities of the wave increase.

Figure 2a shows plots of the reciprocal penetration depths λ_1 and λ_2 corresponding to the components of the wave as functions of the wave frequency and electron temperature. The unshaded region of Fig. 2a is the region where pure surface waves exist; the spatial distribution of the potential in the depth of the plasma for these waves is shown in Fig. 3a. The potential falls off exponentially at large distances from the plasma boundary. The shaded region is the region where generalized surface waves exist. Evidently, the higher the temperature T_e, i.e., the more intense the thermal motion of the plasma electrons is, the more deeply the wave penetrates into the plasma. When the external magnetic field is lowered, the generalized surface wave region vanishes. The spatial distribution of the potential in the depth of the plasma for the generalized surface waves is shown in Fig. 3b. The spatial structure of the potential of the generalized surface waves is such that it oscillates in the depth of the plasma as it falls off exponentially at large distances from the plasma boundary.

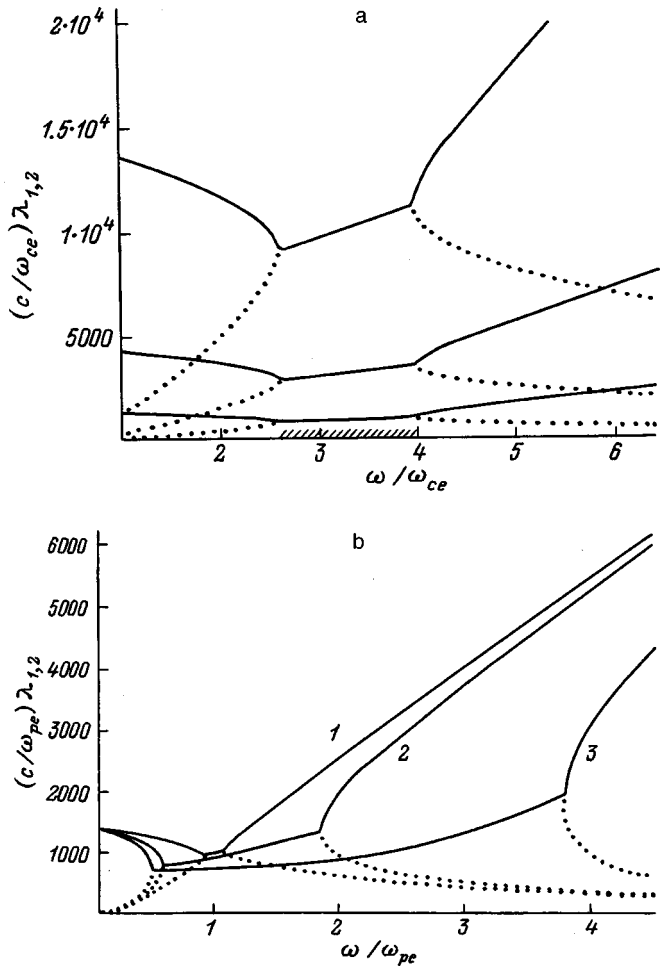

FIG. 2. a: The reciprocals of the penetration depths λ_1 and λ_2 of the corresponding components of the wave into the plasma as functions of the wave frequency for different electron plasma temperatures: $T_e = 100$ (upper pair of curves), 1000 (middle pair of curves), and 10 000 K (bottom pair of curves); these curves were constructed for a plasma density $n_0 = 10^{10}$ cm^{-3} and external magnetic field $H_0 = 100$ G; b: the reciprocals of the penetration depths λ_1 and λ_2 of the corresponding components of the wave into the plasma as functions of the wave frequency for a fixed plasma density $n_0 = 10^{10}$ cm^{-3} and different magnetic fields; the electron temperature is $T_e = 1000$ K. $\omega_{ce} = 0.1\omega_{pe}$ (*1*), $\omega_{ce} = 1\omega_{pe}$ (*2*), $\omega_{ce} = 3\omega_{pe}$ (*3*).

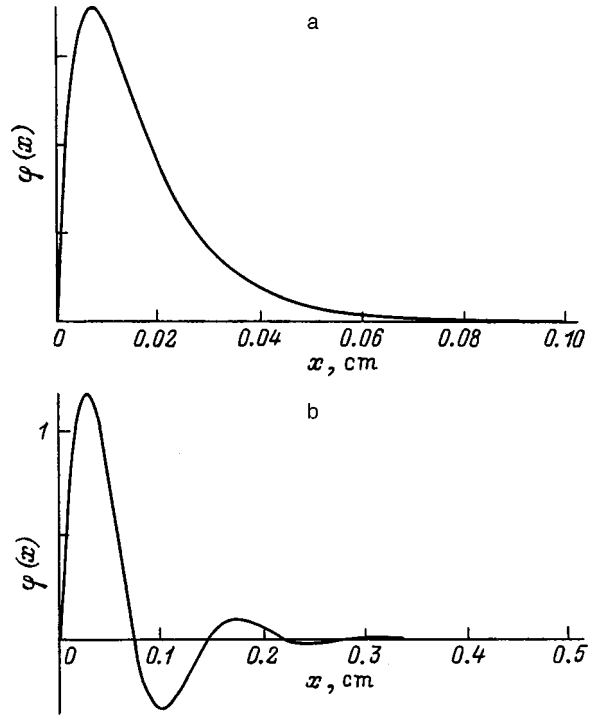

FIG. 3. Spatial distribution of the potential $\varphi(x)$ for pure surface waves ($\omega = 2\omega_{ce}$) (a) and generalized surface waves $\omega = \omega_{pe}$) (b). The plasma temperature is 1000 K, the density is $n_0 = 10^{10}$ cm^{-3}, and the external magnetic field is $H_0 = 100$ G.

Figure 2b shows the reciprocal penetration depths λ_1 and λ_2 of the corresponding wave components as functions of the wave frequency for different values of the parameter $\Omega = \omega_{pe}/\omega_{ce}$ (for a fixed plasma density). A numerical analysis shows that, for any finite external magnetic field H_0, there is a finite range of frequencies corresponding to generalized surface waves. Evidently, the higher the external magnetic field is (i.e., the lower Ω is for given ω_{pe}), the broader the generalized surface wave region will be. It is also apparent that the smaller Ω is, the smaller the penetration depth of the wave into the plasma will be.

The variation in the dispersion curves with $\Omega = \omega_{pe}/\omega_{ce}$ (for fixed plasma density and variable external magnetic field) is illustrated in Fig. 4. Evidently, as the field H_0 is increased (i.e., as Ω is lowered for fixed ω_{pe}), the

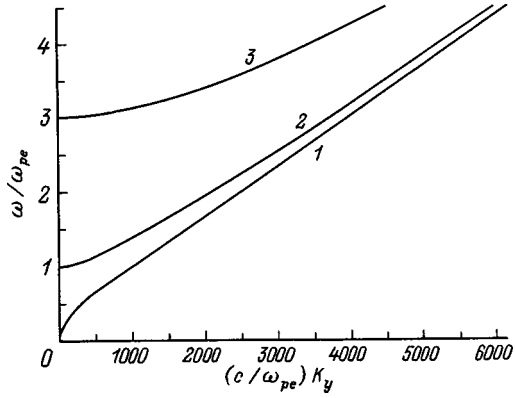

FIG. 4. The dispersion for different values of $\Omega \equiv \omega_{pe}/\omega_{ce}$ (ω_{pe} is fixed, while the external magnetic field and, therefore, ω_{ce} vary). The plasma temperature is $T_e = 1000$ K, the plasma density is $n_0 = 1 \times 10^{10}$ cm^{-3}; $\Omega = 10$ (*1*), 1 (*2*), 0.33 (*3*).

region in which surface waves exist will be shifted toward higher frequencies. This means that the excitation frequencies of the waves can be controlled by varying the external magnetic field.

In this paper we have shown that surface-type waves can propagate along the boundary of a plasma with an ideally conducting metal when there is an external magnetic field perpendicular to the interface. A dispersion relation for these waves and expressions for the penetration depths of the wave field into the plasma have been obtained and these have been studied numerically for different plasma parameters. It has been shown that surface-type waves exist in this geometry for frequencies $\omega > \omega_{ce}$. It has also been shown that in this kind of system with a nonzero external magnetic field, there is a range of frequencies within which the wave is a generalized surface wave. The range of frequencies in which the generalized surface wave exists can be controlled by varying the external magnetic field.

[1] N. A. Azarenkov, A. N. Kondratenko, and K. N. Ostrikov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **36**, No. 5 (1993).
[2] A. N. Kondratenko, *Surface and Volume Waves in Bounded Plasmas* [in Russian], Énergoatomizdat, Moscow (1985), 232 pp.
[3] N. N. Beletskiĭ, A. A. Bulgakov, S. I. Khankina, and V. M. Yakovenko, *Plasma Instabilities and Nonlinear Phenomena in Semiconductors* [in Russian], Naukova Dumka, Kiev (1984), 192 pp.
[4] A. N. Kondratenko and V. M. Kuklin, *Elements of Plasma Electronics* [in Russian], Énergoatomizdat, Moscow (1985), 320 pp.
[5] A. V. Nedospasov and M. Z. Tokar', ''Edge plasmas in tokamaks,'' in *Reviews of Plasma Physics*, Vol. 18, edited by B. B. Kadomtsev [Consultants Bureau, New York (1993); Énergoatomizdat, Moscow (1990), p. 68.].
[6] J. Margot, M. Moisan, and A. Ricard, Appl. Spectrosc. **45**, 260 (1990).
[7] I. Zhelyazkov and V. Atanassov, Phys. Rep. **255**, 79 (1995).
[8] N. A. Azarenkov, I. B. Denisenko, and K. N. Ostrikov, J. Phys. D **28**, 2465 (1995).
[9] N. A. Azarenkov, Zh. Tekh. Fiz. **57**, 1109 (1987) [Sov. Phys. Tech. Phys. **32**, 681 (1987)].

Translated by D. H. McNeill

Mechanisms for formation of the electron distribution function in the positive column of discharges under Langmuir-paradox conditions

A. A. Kudryavtsev and L. D. Tsendin

St. Petersburg State University, 198904 St. Petersburg, Russia;
St. Petersburg State Technical University, 195257 St. Petersburg, Russia
(Submitted November 13, 1998)
Zh. Tekh. Fiz. **69**, 34–41 (November 1999)

The form of the electron distribution function in the positive column of low-pressure discharges is examined under conditions such that the electron mean free path exceeds the vessel radius. Its formation is analyzed taking all major factors into account, including elastic and inelastic collisions, radial and axial electric fields, and the loss of fast electrons to the wall. It is shown that the main mechanism controlling the fast part of the distribution function is the loss of electrons to the wall, which is determined by the scattering of electrons into a comparatively small loss cone that depends on the relationship between the axial and radial components of the velocity. Since the elastic collision rate for all elements has a weak dependence on the energy beyond the ionization threshold, ultimately the high-energy part of the electron energy distribution function in the positive column of low-pressure discharges is nearly Maxwellian. The subthreshold portion of the distribution function, in turn, is determined by the energy diffusion, in a comparatively strong field, of Maxwellian electrons which arrive after inelastic collisions. The final electron distribution function is well approximated by an exponential with a single slope over the entire energy range. Only within a narrow range of scattering angles is the electron distribution function strongly depleted by the loss of electrons to the vessel walls. In the end, it is concluded that this phenomenon, like the Langmuir paradox, may be related to aspects of the physics of the formation of the electron distribution function owing to a combination of already known mechanisms, rather than to a hypothetical mechanism for thermalization of the electrons, as assumed up to now in the literature. A comparison of solutions of the model kinetic equation given here with published Monte Carlo calculations and experimental data shows that they are in good agreement.
 [S1063-7842(99)00711-4]

The electron distribution function is traditionally one of the basic questions of gas discharge physics, since the electrons substantially determine the properties of the plasma as a whole. (For example, see the up to date reviews of work on the kinetics of electrons in gas discharges in Ref. 1.) It has been established that, even in such a comparatively simple object as the positive column of glow discharges, the form of the electron distribution function is determined by a variety of parameters and processes. These, in particular, include self-consistent electric fields, elastic and inelastic electron collisions, spatial transport, etc. In the diffusion regime, where the electron mean free path $\lambda < R$, the vessel radius, collision processes, and the loss of electrons to the walls cause a rapid (exponential) fall-off of the electron distribution function with different slopes in the corresponding energy intervals, which has been confirmed in numerous experiments and calculations. At the same time, one of the most mystifying phenomena in the modern physics of gas discharge plasmas has been observed in low-pressure (below 0.01–0.1 Torr·cm) discharges ever since the pioneering work of Langmuir in the 1920s.[2] The linearity of probe current–voltage characteristics in semilog plots indicated that the electron energy distribution function was close to Maxwellian. The radial gradient of the potential $\phi(r)$ which exists in a plasma traps electrons with energies below $e\phi_w$. (ϕ_w is the potential difference between the axis and wall of the vessel.) Contrary to expectations, however, as the energy increases at $\varepsilon > e\phi_w$ the electron energy distribution function does not fall off faster, even though the corresponding electrons can freely overcome the radial potential difference and, in a short mean free time

$$t_f = R/\sqrt{2\varepsilon/m}, \tag{1}$$

escape to the walls and recombine there.

Langmuir returned repeatedly to this paradoxical discrepancy between experiment and an estimate that seemed based on physically obvious considerations. In this regard, D. Gabor wrote[3] of the ''worst discrepancy known to science.'' The term ''Langmuir paradox,'' itself, belongs to Gabor and has entered the scientific vocabulary. It has been long discussed in the literature.[4–16] Right up to the present time, in the literature primary attention has been directed at searches for a universal mechanism for Maxwellization of the electron distribution function, although these attempts have been, as yet, unsuccessful. Here the Langmuir paradox itself is essentially proclaimed, since no analyses have been made of the formation of the electron distribution function taking all the major factors into account. Thus, the Langmuir

paradox remains a topical and fundamental problem in the physics of gas discharges, which must be solved in order to understand the properties of low-temperature plasmas and to justify the validity of the probe diagnostics used on them.

A Maxwellian distribution usually develops in the course of random interactions (collisions) of particles among themselves or with a thermostat. Since interelectronic collisions are rare under ''Langmuir-paradox'' conditions, various mechanisms for Maxwellization of the electrons have been discussed widely in the literature. Langmuir, himself, advanced the hypothesis that collective interactions of the electrons with turbulent electric fields that develop during plasma (Langmuir) oscillations may play a role.[2] Oscillations in the wall layers of a discharge have been observed[3] through the deflection of an electron beam. Similar observations have been reported elsewhere.[8,9] On the other hand, these oscillations have not been observed in other experiments done specially under similar conditions,[10–12] and doubts have been raised as to whether these oscillations are responsible for Maxwellizing the electron distribution function.[13] It should also be noted that turbulent oscillations are not in thermodynamic equilibrium and cannot, in general, serve as a thermostat. Thus, interactions of the electrons with these oscillations, even if they occur, do not necessarily produce a Maxwellian electron distribution function. Other mechanisms for Maxwellization of the electron distribution function have been discussed,[7] but no satisfactory explanation of the observed effects has yet been obtained.

The proposition can be advanced that the Langmuir paradox is not related to some unknown mechanism for Maxwellization, but to the physical features of the formation of the electron distribution function under these conditions. Under these conditions a combination of already known mechanisms may produce an electron energy distribution function that is close to an exponential with a constant slope. As an example, consider the well-known Druyvesteyn–Davydov distribution, which exists for an elastic balance of the energies of electrons in a collisional plasma (for $\lambda > R$) in an external electric field,[4]

$$f_0(w) = C \exp\left(- \int_0^w \frac{3\,\delta m \nu^2(w)\,dw}{2e^2E^2} \right). \tag{2}$$

It can be seen that for a constant elastic collision rate ($\nu(w) = \text{const}$), it has a Maxwellian form in the absence of collisions among the electrons themselves or with a thermostat.

In this paper we present model calculations of the electron energy distribution function in the positive column of low-pressure glow discharges and verify the possibility of approximating it by an exponential with a single slope over a wide range of energies. Qualitative arguments are presented to the effect that a combination of already known mechanisms operating in low-pressure discharges can create electron energy distribution functions that are close to Maxwellian.

The kinetic equation for the electrons, including elastic and inelastic collisions, radial and axial electric fields, and the loss of electrons to the walls, has been examined previously for the case $\lambda > R$ of interest to us.[17] There the initial kinetic equation for the distribution function was analyzed and simplified on the basis of the following physical considerations. Given the inequality $\nu \gg \nu_i$ between the corresponding collision frequencies in a steady-state glow discharge, an electron which should participate in a single ionization event in its lifetime will undergo many elastic collisions. This inequality is usually satisfied in the range of energies of interest. Thus, because of the large difference between the momentum and energy relaxation times of the electrons, the electron distribution function, as in the case $\lambda < R$, can be represented as the sum of isotropic $f_0(r,\nu)$ and anisotropic $f_1(r,\nu,\theta)$ components. Since the anisotropy of the electron distribution function is caused by two independent factors, drift in the axial field E_z and radial escape of fast electrons to the wall of the vessel, f_1 itself has two components. If the anisotropy is small, then they can be treated independently. For the first component, this is ensured by the smallness of the energy acquired along the mean free path in the field E_z compared to the total energy of the electron. Since it is traditionally discussed in all papers on the kinetics of the positive column, we shall not examine it here. Smallness of the second component, on the other hand, is associated with smallness of the loss cone within which the velocity of an electron leaving the plasma falls.[17] This condition is usually satisfied for most of the electrons of interest, whose energies moderately exceed the wall potential. (See Eq. (6) below.)

Electrons with energies $\varepsilon < e\phi_w$, the potential difference between the axis and wall of the vessel, are trapped in the volume by the ambipolar potential fall $e\phi_a$ and the wall potential jump $e\phi_s$, so that $\phi_w = \phi_a + \phi_s$. Thus, the main part of the distribution function for the trapped electrons, f_0, depends only on the total energy[17]

$$\varepsilon = w(r) + e\phi(r) \tag{3}$$

(the kinetic $w = m\nu^2/2$ plus potential $e\phi(r)$ energies).

Electrons with energies $\varepsilon > e\phi_w$ can escape to the wall. Under low-pressure discharge conditions, the anisotropic component of the electron distribution function associated with the escape of fast electrons can be substantial. An electron leaves the volume if it falls within a loss cone $\delta\Omega$ in which the normal component of its energy exceeds the existing potential barrier.[17] If we denote the polar angle between the velocity vector of an electron and the radius vector r by $\alpha(\varepsilon,\mu,r)$ (μ is the angular momentum of the electron), then $\delta\Omega$ can be written in the form[18,19]

$$\delta\Omega = 2\pi(1 - \cos\alpha(\varepsilon,\mu,r)). \tag{4}$$

In order to escape to the wall, an electron must be elastically scattered into this loss cone, which is equivalent to loss of electrons at a rate

$$\nu_w = 2\nu(\delta\Omega/4\pi). \tag{5}$$

The coefficient 2 shows up because of the symmetry of scattering at angles α and $\pi - \alpha$. As can be seen from Eq. (5), the electrons are lost much more slowly than through free flight to the walls, which takes a time t_f [Eq. (1)]. For a plane geometry,[18]

$$\cos\alpha(\varepsilon,r)=\sqrt{\frac{e\phi_w-e\phi(r)}{\varepsilon-e\phi(r)}}. \tag{6}$$

In a cylindrical geometry, electrons with total energy $\varepsilon>e\phi_w$ can escape to the wall if their perpendicular energy ε_p exceeds the sum of the potential and centrifugal energies at the wall, i.e.,[17,19]

$$\varepsilon_p=mv_r^2/2+\mu^2/2mr^2+e\phi(r)\geqslant e\phi(R)+\mu^2/(2mR^2), \tag{7}$$

where $mv_r^2/2=(\varepsilon-e\phi(r))\cos^2\alpha$.

As long as the loss cone is small ($\delta\Omega\ll 4\pi$), for isotropic elastic scattering of the electrons, their distribution function outside this cone, f_0, can also be regarded as isotropic and dependent only on ε.[17]

The character of the anisotropy in the electron distribution function varies substantially, depending on whether it is associated with an axial field or with loss to the wall for isotropic scattering. In the first case, the electron distribution function differs little from isotropic at all angles. Mathematically, this is reflected in the fact that the coefficients in the ordinary expansion of the electron distribution function in terms of Legendre polynomials fall off rapidly with the harmonic number. In the second case, however, the electron distribution function is almost isotropic outside a small loss cone and practically zero inside it. In other words, the electron distribution function has a discontinuity at its boundary. Outside the small cone, it is almost identical to the isotropic part of the electron distribution function, f_0. Within the loss cone, the distribution is also independent of angle. Therefore, all the coefficients in the expansion of the anisotropic part of the electron distribution function in terms of spherical harmonics are also small (proportional to $\delta\Omega$), but do not fall off with the harmonic number. Thus, it is inappropriate to use this expansion and it is more convenient to consider $f_0(\varepsilon)$ and the electron distribution function inside the yield cone, $F(\varepsilon,r,\delta\Omega)$, separately.

After averaging over the volume, the kinetic equation for $f_0(\varepsilon)$ can be written in the form[17]

$$\frac{\partial}{\partial\varepsilon}\overline{D}_E\frac{\partial f_0(\varepsilon)}{\partial\varepsilon}=\overline{\nu}_{\rm ex}(\varepsilon)f_0(\varepsilon)-\overline{\nu}_{\rm ex}(\varepsilon+\varepsilon_{\rm ex})f_0(\varepsilon+\varepsilon_{\rm ex})+\overline{\nu}_i(\varepsilon)f_0(\varepsilon)-2\overline{\nu}_i(2\varepsilon+\varepsilon_i)f_0(2\varepsilon+\varepsilon_i)+\overline{\nu}_w f_0(\varepsilon), \tag{8}$$

where the coefficients of Eq. (8), obtained by averaging over the radius, are

$$\overline{D}_E=\frac{2e^2E_z^2}{3m}\int_0^{r_1\varepsilon}\frac{(\varepsilon-e\phi(r))^{3/2}rdr}{\nu(\varepsilon-e\phi(r))}, \tag{9}$$

$$\overline{\nu}_j(\varepsilon)=\int_0^{r_j}\nu_j(\varepsilon-e\phi(r))(\varepsilon-e\phi(r))^{3/2}rdr. \tag{10}$$

The radius $r_1(\varepsilon)$ determines the point at which an electron turns and is found from the condition $\varepsilon=e\phi(r_1(\varepsilon))$, while $r_j(\varepsilon)$ is found from the condition $e\phi(r_j(\varepsilon))=\varepsilon-\varepsilon_j$, where the index $j=ex, i$, or w.

In Eqs. (9) and (10), E_z is the external electric field and $\nu_{\rm ex}$ and ν_i, respectively, are the rates of the inelastic processes for excitation and ionization. For simplicity, it is assumed that there is only one excitation process with a threshold $\varepsilon_{\rm ex}$ and only direct ionization from the ground state. For concreteness, the collision integral for ionization is written assuming that the kinetic energy is divided evenly between the incident and product electrons.[18]

The loss cone $\delta\Omega$ is essentially empty, since the corresponding electrons escape instantaneously to the wall (over at time t_f [Eq. (1)] and, neglecting the effect of the radial electric field in the plasma on the fast electrons, the kinetic equation for the distribution function $F(\varepsilon,r,\Omega)$ inside this cone can be written in the form

$$\frac{\partial F(\varepsilon,r,\delta\Omega)}{\partial r}=\frac{1}{4\pi\lambda}f_0(\varepsilon), \tag{11}$$

where $\lambda=v/\nu$ is the electron mean free path. Evidently, $F\ll f_0$.

The self-consistent wall potential $\varphi_w=\phi_a+\phi_s$ is found by equating the number of ionizations Z_i to the number of electrons escaping to the wall, W_e, i.e.,

$$Z_i=\frac{2}{R^2}\int_0^R rdr\int_{\varepsilon_i}^{\infty}\nu_i(w)\sqrt{w}\,f_0(\varepsilon)d\varepsilon=W_e$$

$$=\frac{2}{R^2}\int_0^R rdr\int_{e\phi_w}^{\infty}\nu_w(w)\sqrt{w}\,f_0(\varepsilon)d\varepsilon. \tag{12}$$

Since $\nu\gg\nu_i$, it is evident from Eq. (12) that the wall potential $e\phi_w$ exceeds the ionization potential ε_i and the major portion of the electrons are trapped in the volume.

The axial electric field is found from the condition that the discharge is stationary, i.e., by equating the average number of ionizations per unit volume to the number of ions lost to the wall, W_i,

$$Z_i=W_i=\overline{n}_e/\tau_i, \tag{13}$$

where $\overline{n}_e$ is the volume averaged electron density and τ_i is the ion lifetime.

Since the ion mean free path usually obeys $\lambda_i<\lambda$, different relationships between λ_i and R are possible under the present conditions ($\lambda>R$). In general, τ_i is conveniently written in the form

$$\tau_i=\Lambda^2/D_s, \tag{14}$$

and for the effective coefficient D_s one can use an interpolation,[17] which transforms from the collisional Schottky regime to the collisionless Tonks-Langmuir regime in the limits,

$$D_s=D_a/(1+2Z_i/(\overline{n}_e\nu_{ia})), \tag{15}$$

where $D_a=D_i(1+T_e/T)$ is the customary ambipolar diffusion coefficient and $\nu_{ia}=Nv_i\sigma_{ia}$ is the ion-neutral (atom) collision rate.

In order to verify the accuracy of the approximate kinetic equation (8), it is appropriate to compare it with the results of an independent Monte Carlo method. For the conditions of interest to us, the most detailed work in this re-

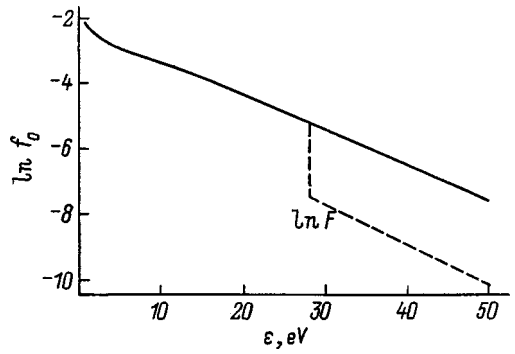

FIG. 1. $f_0(\varepsilon)$ calculated according to Eq. (7). $N=10^{14}\,\mathrm{cm}^{-3}$, the vessel radius $R=1$ cm, $\phi_w=28.5$ V, and $E_z=1$ V/cm.

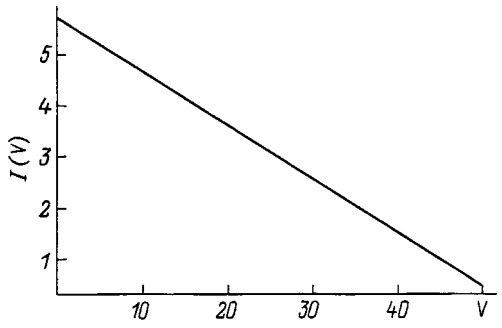

FIG. 2. The probe current I as a function of the potential V. The conditions are the same as in Fig. 1.

spect is Ref. 19, in which the electron distribution function in the positive column of a glow discharge for an argon-like gas was modelled over a wide range of pressures (3 mTorr to 3 Torr). The authors were mainly guided by the collisional case ($\lambda > R$), for which they observed a change in the slope of the electron energy distribution function for $\varepsilon > \varepsilon_{\mathrm{ex}}$ owing to inelastic processes and the loss of electrons to the wall. A comparative analysis of these data[19] shows that when the pressure is lowered ($p < 10$ mTorr), the relative contribution of wall losses increases and there is a clear tendency for the electron energy distribution function to flatten out to an exponential with a single slope throughout the entire energy range.

In the following comparisons we shall use the cross sections for argon, as in Ref. 19:

$$\sigma = 1.59\times 10^{-15}\begin{cases} w/\varepsilon_{\mathrm{ex}}, & w<\varepsilon_{\mathrm{ex}} \\ \sqrt{\varepsilon_{\mathrm{ex}}/w}, & w>\varepsilon_{\mathrm{ex}} \end{cases}, \mathrm{cm}^2, \tag{16}$$

$$\sigma_{\mathrm{ex}} = 1.56\times 10^{-16}\ln(w/\varepsilon_{\mathrm{ex}}),\ \mathrm{cm}^2, \tag{17}$$

$$\sigma_i = 3.18\times 10^{-16}\ln(w/\varepsilon_i)/(w/\varepsilon_i),\ \mathrm{cm}^2, \tag{18}$$

$$\sigma_{ia} = 4\times 10^{-15},\ \mathrm{cm}^2. \tag{19}$$

On averaging in accordance with Eqs. (9) and (10), the profile of the ambipolar potential can be approximated by the parabola

$$e\phi(r) = e\phi_a(r/R)^2. \tag{20}$$

It is clear from Eqs. (16) and (19) that in the case $\lambda > R$ of interest to us, the ions are also in a drift regime. Thus, in the following calculations we have followed Tonks and Langmuir and taken $e\phi_a = 1.145T_e$ in Eq. (20) and

$$\tau_i = R/v_{is} \tag{21}$$

in Eq. (13), where $v_{is} = \sqrt{T_e/M}$ is the ion sound velocity.

The model calculations in Ref. 19 showed that the simple expression (6) for $\delta\Omega$ in the case of a plane geometry is a good approximation for a cylinder, as well.

As a comparison, Fig. 1 shows f_0 calculated according to Eq. (7) for the conditions of Ref. 19 (neutral gas density $N=10^{14}\,\mathrm{cm}^{-3}$ and vessel radius $R=1$ cm). The self-consistent values of the wall potential $\phi_w = 28.5$ eV and longitudinal field $E_z = 1$ V/cm turned out to be close to those calculated in Ref. 19 (30 eV and 1.02 V/cm, respectively). A semilog plot of the electron energy distribution function is clearly close to linear with an effective electron temperature $T_e = 9.3$ eV. Significant deviations from a Maxwellian distribution function are observed only at low energies. The calculations also confirm the well known fact from the literature[14] that T_e in low-pressure discharges is set near the threshold $\varepsilon_{\mathrm{ex}}$ for inelastic processes. Physically, such high temperatures are necessary in order for an electron to complete an ionization event during the short ion mean free time τ_i to reach the wall (21). Since in this model we are only considering direct ionization from the ground state, Eqs. (12) and (13) are independent of the absolute electron density and the electron energy distribution function is expressed in arbitrary units. If we use the electron energy distribution function shown in Fig. 1 to construct a plot of the probe current

$$I(eV) = \int_{eV}^{\infty} (\varepsilon - eV)\, f_0(\varepsilon)\, d\varepsilon, \tag{22}$$

from which the Langmuir paradox was first observed in the literature, then we obtain a straight line which is essentially identical to an exponential (Fig. 2). This is because an integral dependence of the type (22) has a weaker dependence on the details of the behavior of the electron energy distribution function. It is known that the finite magnitude of the differentiated signal d at the probe means that the electron energy distribution function is significantly distorted at energies near the plasma potential for small potentials of order d, which are around $(0.3-1)T_e$ in the experiment. Thus, in the actually measured second derivatives of the probe current $I''(eV)$, the low energy part of the electron energy distribution function is not recorded. As an example, Fig. 3 also

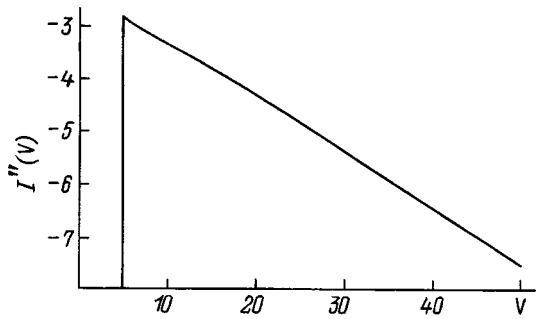

FIG. 3. The second derivative $I''(V)$ of the probe current for a differentiation step size $d=5$ V. Conditions as in Fig. 1.

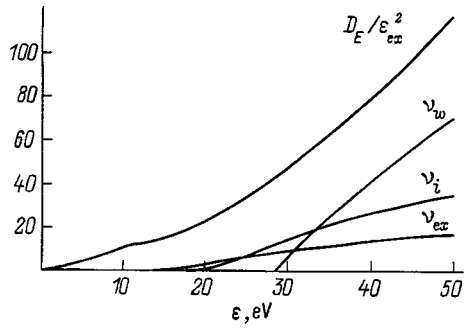

FIG. 4. The terms in the kinetic equation (8) for the conditions of Fig. 1.

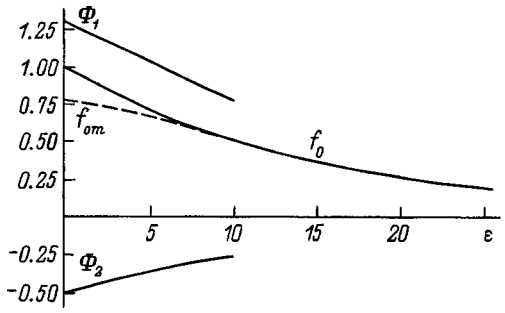

FIG. 5. The components of the electron energy distribution function (27) for $T_e=\varepsilon_{\rm ex}$.

shows a model calculation of the second derivative of the probe current with a finite differentiation step $d=5$,

$$I''(V)=\frac{I(V+d)+I(V-d)-2\cdot I(V)}{d^2}. \tag{23}$$

Evidently, for $\varepsilon>5$ eV, $I''(eV)$ is close to exponential and coincides with the electron energy distribution function of Fig. 1. The calculations also show that curves analogous to those of Figs. 1 and 2 are obtained for pure gases when $\lambda>R$ under all these conditions. In order to understand the reasons for such surprising behavior of f_0, we turn to Fig. 4, which shows the terms in the kinetic equation (8) for the conditions of Fig. 1. Obviously, for $\varepsilon>e\phi_w$ the rate of loss to the wall significantly exceeds the inelastic collision rates, while the energy dependences of $\overline{D}_E(\varepsilon)$ and $\overline{\nu}_w(\varepsilon)$ are similar. Here the kinetic equation (8) for the fast part of the electron energy distribution function with $\varepsilon>e\phi_w$ reduces to

$$\frac{\partial}{\partial\varepsilon}\,\overline{D}_E\,\frac{\partial f_0(\varepsilon)}{\partial\varepsilon}=\overline{\nu}_w\,f_0(\varepsilon), \tag{24}$$

and, roughly speaking, the electron ''temperature'' for the fast part of the electron energy distribution function,

$$T_w=\sqrt{\overline{D}_E/\overline{\nu}_w}, \tag{25}$$

is practically constant and independent of the energy. The reason for this is the weak, decreasing energy dependence of the elastic cross section (16) at high energies. This leads to a small energy dependence for $\nu(w)$ and, therefore, for T_w in Eq. (25). It is well known from the literature[20] that at high energies $\varepsilon\geqslant\varepsilon_i$ the elastic scattering cross sections fall off slowly with increasing energy and behave this way for almost all elements. Thus, this sort of dependence is quite universal. At high energies, therefore, T_w [Eq. (25)], is nearly constant for essentially any pure gas and the electron energy distribution function for $\varepsilon>e\phi_w$ should be close to Maxwellian,

$$f_0(\varepsilon)\approx\exp(-\varepsilon/T_w). \tag{26}$$

We now consider some possible reasons why the electron energy distribution function may have the same slope in the elastic energy range, $\varepsilon<\varepsilon_{\rm ex}$, as well. At first glance, this seems paradoxical, since it is well known[20] that the actual energy dependences of $\nu(w)$ are different in this region for different gases. Let us write down a model solution of the kinetic equation (8) for $\varepsilon<\varepsilon_{\rm ex}$ with the fast part of the electron energy distribution function in the form (26),

$$\begin{aligned}f_{0m}(\varepsilon)=\Phi_1-\Phi_2=&(\exp(-\varepsilon_{\rm ex}/T_w)\\&+\exp(-2\varepsilon_{\rm ex}/T_w))(1+(\varepsilon_{\rm ex}-\varepsilon)/T_w)\\&-\exp(-(\varepsilon+\varepsilon_{\rm ex})/T_w),\end{aligned} \tag{27}$$

which represents the difference between the linear Φ_1 and exponential Φ_2 functions. In the elastic energy range, the model electron energy distribution function (27) is evidently never Maxwellian, even for a Maxwellian fast part (26). The functions (26) and (27) are close to one another for $\varepsilon\sim\varepsilon_{\rm ex}$ and diverge with decreasing energy. Since their differences are greatest at low energies, for $\varepsilon/T_e\ll1$, on expanding the exponent in Φ_2 in a Taylor series, near zero we obtain

$$f_0(\varepsilon\approx0)\approx(\varepsilon_{\rm ex}/T_w)\exp(-\varepsilon_{\rm ex}/T_w). \tag{28}$$

The expression on the right of Eq. (28) is close to unity [i.e., to the corresponding values of Eq. (26)] only for $\varepsilon_{\rm ex}\approx T_w$ and falls off sharply with decreasing T_w. Thus, the closeness of the approximate electron energy distribution function in the elastic energy range to an exponential with the temperature of the fast portion of the electron distribution is determined by the parameter $\varepsilon_{\rm ex}/T_w$. This sort of approximation is possible only for high $T_w\approx\varepsilon_{\rm ex}$. This is illustrated in Fig. 5, which shows the components of the model electron energy distribution function according to Eqs. (26) and (27) for $\varepsilon_{\rm ex}/T_e=1$, and in Fig. 6, which shows the model electron energy distribution functions for $\varepsilon_{\rm ex}/T_e=0.5$, 1, and 5. Evidently, the complete electron energy distribution function should never be strictly Maxwellian. It can be approximated

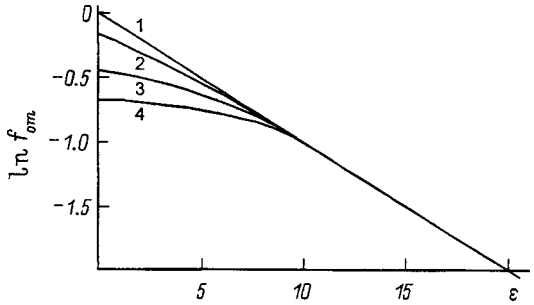

FIG. 6. Model electron energy distribution functions (27) for $\varepsilon_{\rm ex}/T_e=5$ (*2*), 1 (*3*), and 0.5 (*4*) and the electron energy distribution function of Eq. (26) (*1*).

by an exponential with a single slope only when the electron temperature is high and close to the inelastic threshold.

Therefore, in order for the entire electron energy distribution function in low-pressure discharges to be approximated by a Maxwellian distribution, two major conditions must be satisfied: the elastic cross section must fall off slowly at high energies (beyond the ionization potential of the working gas) and the electron temperature must be high (on the order of the inelastic threshold). Failure of either of these conditions will cause the electron energy distribution function to deviate significantly from an exponential with a single slope, i.e., the Langmuir paradox will be absent. In order to establish whether this much discussed phenomenon occurs or not, that is, to determine the unknown mechanism for mawellianization of the distribution and quantitatively determine its efficiency under various conditions, or confirm its absence, it is necessary to measure and calculate the electron energy distribution function in detail, along with the radial electric field profile and the potential jump at the wall, under identical conditions. Unfortunately, essentially no complete sets of these measurements have been published. At present, only we are doing this kind of work. Thus, based on an analysis of the known experimental data of various authors, in this paper we shall present qualitative evidence that distributions close to Maxwellian can, in principle, develop without relying on any additional mechanisms for Maxwellization.

Kagan and colleagues at St. Petersburg State University made systematic measurements of the electron energy distribution function in the positive column of discharges in various gases: mercury, cadmium, helium, neon, argon, and mixtures of these at various pressures and currents. Their results were summarized in a review[6] where the following features were noted. Maxwellian electron energy distribution functions were observed in all the single component systems. In mercury and the heavy inert gases, this corresponded to pressures of up to hundredths of a Torr, and in helium and neon up to a few tenths of a Torr; that is, the upper bound on the pressure increased as the corresponding elastic scattering cross sections decreased. As the pressure was raised further (roughly after attaining the condition $\lambda < R$), the electron energy distribution function was depleted in the elastic energy range, a fact that was attributed[6] to an increased relative contribution from inelastic excitation and ionization processes. (See Figs. 7 and 13–17 of Ref. 6.)

Another series of experiments was carried out at a fixed low pressure, 3 mTorr, in mercury (i.e., under conditions such that the Langmuir paradox is observed) and with added inert gases (He, Ne, Ar, Kr, and Xe). Apparently, here the electron distribution was only changed by the buffer gas. Nevertheless, depletion of the electron energy distribution function beyond the threshold for inelastic collisions in mercury was observed in the experiments. This was rather surprising, since in this energy range only the rate of elastic collisions with the inert gas atoms increased with the mercury pressure held fixed. (See Figs. 9–12 of Ref. 6.) Since adding a buffer gas in these amounts could only change the time for loss to the walls, it was suggested[5,6] that the role of ''wall'' effects had decreased. This was interpreted as indirect evidence in favor of some unknown ''wall'' mechanism for formation of a Maxwellian electron energy distribution function.

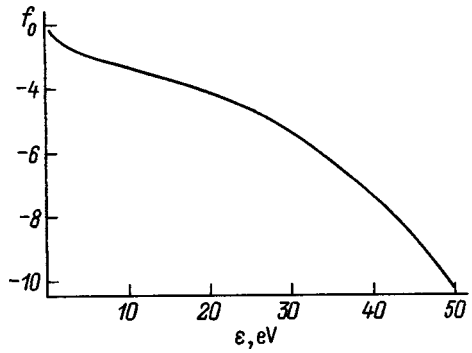

FIG. 7. $f_0(\varepsilon)$ calculated according to Eq. (7) with an increasing elastic scattering cross section, per Eq. (29). The remaining conditions are as in Fig. 1.

These observations[6] can, on the other hand, be explained on the basis of our previous discussion, without invoking any sort of hypothetical mechanisms. In fact, on going to a mixture of gases, when ionization is still proceeding through an easily ionized impurity, elastic scattering of electrons at $\varepsilon > e\phi_w$ can still be controlled by a buffer gas with a sufficiently high ionization potential. In the energy range of interest to us (which is determined by the excitation and ionization of mercury), the elastic scattering cross section, which, according to Eq. (5), determines the electron losses to the walls, can depend strongly on the energy. Here the wall temperature T_w (25) will not be constant and the slope of the electron energy distribution function must change with energy. As an illustration, Figs. 7 and 8 show curves analogous to those in Figs. 1 and 4, but with a rising elastic cross section (16) throughout the energy range,

$$\sigma = 1.59 \times 10^{-15} (w/\varepsilon_{ex})^{1/2}. \tag{29}$$

It is apparent that, in this case, the functions $\overline{D}_E(\varepsilon)$ and $\overline{\nu}_w(\varepsilon)$ are no longer similar to one another and the electron energy distribution functions differ greatly from exponential with a single slope. Evidently, under these experimental conditions the Langmuir paradox should not occur. These effects are entirely consistent with the data of Ref. 6: when 6 mTorr argon was added to 3 mTorr mercury, i.e., the electrons were still in a drift regime, a deviation from a Maxwellian electron

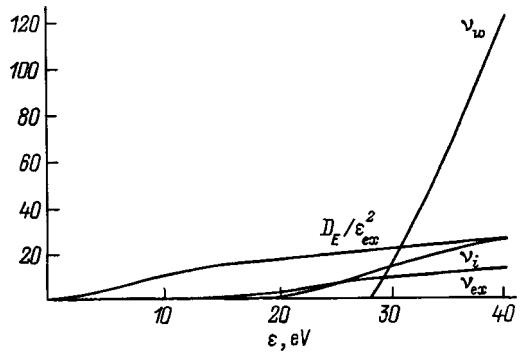

FIG. 8. The terms in the kinetic equation (8) for the conditions of Fig. 7.

energy distribution function could already be observed. (See Fig. 11 of Ref. 6.) Argon has a rapidly rising cross section within this energy range; thus, the picture is similar to that obtained above in Fig. 7. At the same time, when helium, whose elastic cross section varies slowly with energy, was added, nearly Maxwellian electron energy distribution functions were observed[6] up to helium pressures of tenths of a Torr, i.e., up to roughly the same pressures as for discharges in pure helium. (See Figs. 10, 15, and 16 of Ref. 6.)

In Fig. 1 the dashed curve shows $F(\varepsilon, r, \delta\Omega)$ according to Eq. (11). As is to be expected, $F \ll f_0$. In practice, in order to measure the electron distribution function in the loss cone, it is necessary to measure the electron distribution function with angular resolution. When traditional probe methods are used, the measured information on the electron distribution function is, in one way or other, averaged over the angular variable. In particular, a cylindrical probe measures only the distribution with respect to the velocity (energy) perpendicular to its axis. In this situation, the above noted difference in the angular dependences may be observed only when a simultaneous comparison is made of measurements with the probe oriented at different angles to the discharge axis. These features have been observed[14] in a study of the positive column of a mercury discharge at $p = 1.3$ mTorr with a current $I = 80$ mA. Measurements with spherical and cylindrical probes positioned perpendicular to the axis of the vessel (Fig. 3 of Ref. 14), which do not ''sense'' the loss cone, indicate that the electron energy distribution function is Maxwellian over a wide range of energies. Measurements with a cylindrical probe parallel to the axis, however, revealed severe depletion in the second derivative of the probe current at energies beyond the wall potential (Fig. 4 of Ref. 14). This is consistent with the fact that for $\varepsilon > e\phi_w$ this probe only detects electrons in the loss cone (to the walls). Depletion associated with losses to the wall showed up most distinctly in measurements with a flat wall probe (Fig. 3 of Ref. 14). Differences between the second derivatives recorded with a cylindrical probe oriented along and perpendicular to the discharge axis have been noticed elsewhere.[21] For detailed measurement of the angular dependence of the electron distribution function it is appropriate to make the measurements with flat probes at different orientations relative to the discharge axis. This should offer the possibility of detecting an empty loss cone for electrons moving to the vessel walls.

In order to demonstrate the dominant role of electron losses to the wall in the formation of the electron distribution function in a practical way, we can propose changing the boundary conditions at the wall experimentally, for example, by forcibly changing the wall potential. This is most simply done if part of the wall is a conductor (e.g., in the form of a cylindrical metallic insert). When a specified potential that reduces ϕ_w is applied, changes should show up in the recorded current–voltage characteristics or the second derivative of the probe current. Figure 9 shows the results of model calculations for the conditions of Fig. 1 as ϕ_w is varied. It is evident that, as the wall potential is reduced, the depletion of the fast part of the electron energy distribution function owing to changes in T_w begins to show up for small ε. Note that

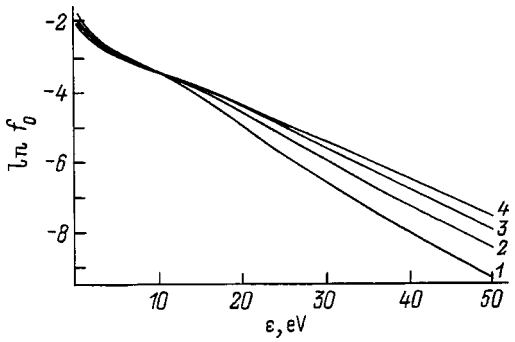

FIG. 9. $f_0(\varepsilon)$ calculated according to Eq. (7) for wall potentials $\phi_w = 13$ (*1*), 18 (*2*), 23 (*3*), and 28.5 V (*4*). The other conditions are as in Fig. 1.

an attempt to do this sort of experiment in a tube with a metal wall has been made before.[22] Since the same Maxwellian electron energy distribution function was observed in a tube with isolated metal walls as in a glass tube, it was concluded that the Langmuir paradox does not depend on the material of the discharge vessel walls. When a positive voltage was applied to the conducting wall of the tube, depletion of the fast portion of the electron energy distribution function was observed and this was explained[22] by a reduced contribution from some unknown ''wall'' mechanism for Maxwellization. As can be seen from Fig. 9, these phenomena are possible without any such hypothetical mechanism.

We have, therefore, shown that in low-pressure discharges where the electron mean free path exceeds the vessel radius, electron loss to the walls is determined by elastic scattering into a narrow loss cone. For the high longitudinal fields which exist under these conditions, the electron energy distribution function can be approximated by an exponential function with a single slope over the entire energy range. We conclude that this effect can, in principle, explain the widely discussed Langmuir paradox.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-18417).

[1] U. Kortshagen and L. Tsendin (Eds.), *Electron Kinetics and Applications of Glow Discharges*, NATO ASI Series B. Physics, Vol. 367, Plenum Press, New York and London (1998).
[2] I. Langmuir, Phys. Rev. **26**, 585 (1925).
[3] D. Gabor, E. A. Ash, and D. Dracott, Nature (London) **176**, 916 (1955).
[4] V. L. Granovskii, *Electrical Currents in Gases* [in Russian], Nauka, Moscow (1971), 272 pp.
[5] Yu. Benke, Yu. M. Kagan, and V. M. Milenin, Zh. Tekh. Fiz. **38**, 1197 (1968) [Sov. Phys. Tech. Phys. **13**, 989 (1968)].
[6] Yu. M. Kagan, in *Spectroscopy of Gas Discharge Plasmas* [in Russian], Nauka, Leningrad (1970), pp. 201–223.
[7] V. M. Milenin and N. A. Timofeev, Zh. Tekh. Fiz. **48**, 1841 (1978) [Sov. Phys. Tech. Phys. **23**, 1048 (1978)].
[8] R. Rompe, S. Ullrich, and H. Wolf, Beitr. Plasmaphys. **4**, 245 (1961).
[9] E. N. Hirsch, Int. J. Electron. **19**, 537 (1966).
[10] G. Gierke, W. Ott, and F. Schwirzke, in *Proceedings of the Sixth International Conference on Phenomena in Ionized Gases (ICPIG)*, Munich (1961), Vol. II, p. 1412.
[11] W. Ott, ''Ein Versuch zur Klarung des Langmuir-Paradoxus,'' Preprint No. 2/19, Inst. Plasmaphys., Munich, (1963), 27 pp.
[12] R. Harp and G. S. Kino, in *Proceedings of the Eighth International Conference on Phenomena in Ionized Gases (ICPIG)*, Paris (1963), Vol. 3, p. 45.
[13] F. W. Crawford and S. A. Self, Int. J. Electron. **18**, 569 (1965).

[14] S. W. Rayment and N. D. Twiddy, Proc. R. Soc. London, Ser. A **340**, 87 (1968).
[15] F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, 2nd ed. [Plenum Press, New York (1984); Mir, Moscow (1987), p. 73].
[16] B. V. Kadomtsev, in M. A. Leontovich (Ed.), *Reviews of Plasma Physics* [in Russian], Vol. 4, Atomizdat, Moscow (1964), p. 325.
[17] L. D. Tsendin and Yu. B. Golubovskiĭ, Zh. Tekh. Fiz. **47**, 1839 (1977) [Sov. Phys. Tech. Phys. **22**, 1066 (1977)].
[18] U. Kortshagen, I. Pukropski, and L. D. Tsendin, Photonics Spectra **51**, 6063 (1995).
[19] U. Kortshagen, G. J. Parker, and J. E. Lawler, Phys. Rev. E **54**, 6746 (1996).
[20] E. W. McDaniel, *Collision Phenomena in Ionized Gases* [Wiley, New York (1964); Mir, Moscow (1967)].
[21] V. A. Godyak, A. A. Kozovnikov, and M. A. Khadar, Vestn. Mosk. Univ. Fiz., Astron. No. 3, 336 (1971).
[22] V. M. Milenin, Candidate's Dissertation [in Russian], Leningrad (1965).

Translated by D. H. McNeill

Effect of the gas dynamic structure of a flow on the parameters of a self-sustained discharge. I

G. A. Baranov and S. A. Smirnov

D. V. Efremov Scientific-Research Institute of Electrophysical Apparatus, 189631 St. Petersburg, Russia
(Submitted October 15, 1997; resubmitted April 13, 1998)
Zh. Tekh. Fiz. **69**, 42–48 (November 1999)

A study is made of self-sustained glow discharges in transverse gas flows and jets. The distributions of the discharge current and voltage over the elements of a sectioned cathode array are measured along with the temperature of the cathode array. The limiting current and discharge voltage corresponding to the transition from a uniformly burning discharge to a contracted state are measured. Two-dimensional and one-dimensional systems of equations for the gas dynamics and vibrational kinetics are used for a numerical analysis of the experimental data, and the results are used to determine the character of the distribution of E/N in the discharge, where E is the electric field and N is the molecular density. The heat balance of the cathode array is calculated. A model is proposed for self-consistently calculating the parameters of the gas flow, the distribution of the current over the cathode array, and the discharge voltage, as well as the values of the ballast resistances. © *1999 American Institute of Physics.* [S1063-7842(99)00811-9]

INTRODUCTION

Studies of the gas dynamic structure of flows in discharge regions are of great importance for developing flow-through gas discharge lasers. They are needed for an adequate calculation of the output characteristics, the degree of optical homogeneity of the active medium, the thermal operating regime of the cathode array, and the limiting characteristics of the discharges. Studies of the effect of the flow parameters on the discharge stability are of decisive priority, since disruption of uniform discharge operation leads to cutoff of the laser emission and to critical overheating of the elements in the cathode array.

The effect of the degree and scale of turbulence, as well of the average flow characteristics on discharge stability have been examined.[1–3] Interpretations of experimental data dealing with this problem generally assume that the overheating-ionization instability is the most to develop. This type of instability develops when the inhomogeneity in the gas density distribution within the discharge volume reaches a critical level.

A local reduction in the density causes E/N to increase and this is followed by a nonlinear rise in the ionization rate. This implies that reducing the inhomogeneity in the gas density and speeding up diffusion processes should improve the limiting characteristics of the discharge. Qualitatively, this is consistent with experiments in which a discharge was stabilized by enhancing the uniformity of the averaged flow parameters and generating small scale turbulence. This approach to interpreting the experiments was not reinforced by an adequate mathematical model which might have been used to calculate the limiting discharge current. The problem arises, first of all, because of the uncertainty in estimating the sites where a critical inhomogeneity develops and its trajectory, and in determining its physical parameters. For example, the spatial scale length of the local perturbation in the density was estimated either from the characteristic size of the nonuniformity in the averaged velocity profile or from the size of the turbulizer or another element of the discharge chamber. In the case of a transverse discharge, the strong inhomogeneities in the velocity profile within the boundary layer and its complicated vortical structure (in the case of a turbulent boundary layer) were neglected.

In this paper we develop an algorithm for an alternative interpretation of the experimental data which is based on an exact quantitative determination of the zone with an anomalously high energy input. Observations of the development of discharge instabilities show that, as a rule, an instability originates near the cathode. The distribution of the discharge current over the cathode elements depends on the distribution of the potentials and, therefore, on the distribution of the gas density in the discharge region. This means that the theoretical study of the discharge stability requires development of an algorithm for calculating the distribution of the discharge current over the regions of the cathode falls. This algorithm will also find applications in numerical modeling of a range of other processes in transverse-discharge lasers. Up to now, the equations for the gas dynamics, vibrational kinetics, and radiation field in a cavity have been solved with distributions of the energy input determined from experimental data.[4]

We have studied the interrelation among the flow parameters, the distributions of the discharge voltage and current, and the values of the ballast resistances over the cathode elements of a sectioned cathode array. The distributions of the discharge voltage and current were measured in a transverse flow and in a highly nonuniform flow. The latter consisted of a system of jets. Gas was blown in through the cathode array with initial velocities near sonic. The numeri-

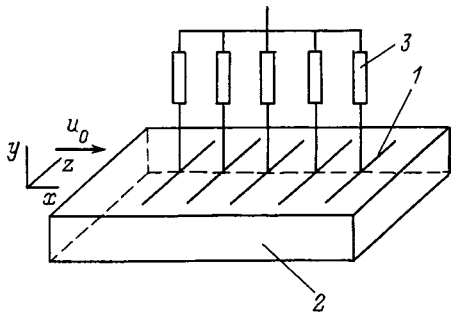

FIG. 1. Sketch of the discharge chamber: *1* — cathode, *2*—anode, *3* — ballast resistance.

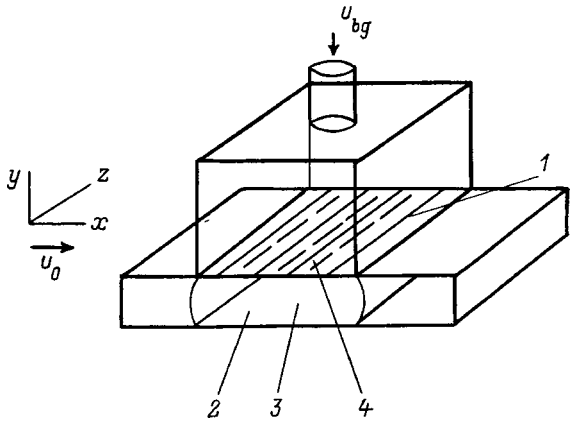

FIG. 2. Sketch of the discharge chamber with a gas inflow: *1* — gap, *2* — anode, *3* — discharge region, *4* — cathode element.

cal analysis of the experimental data employed equations for the flow of a gas in vibrational disequilibrium. The thermal regime of the cathode array was calculated in order to set the temperature boundary conditions. These calculations were tested against experimental data.

THE TEST STAND

The experiments to determine the value of E/N averaged over the channel height were done on an air flow in an open test stand[5] in two regimes: with tranditional transverse air flow through the discharge region and by feeding air into the discharge region only through slits in the cathode array (i.e., in the latter case without an initial transverse flow at the inlet to the discharge region). The distance between the converging channel attached to the channel for the purpose of equilibrating the velocity profile and the first cathode was 1.2 m. Over the section from 0.9 to 1.2 m the channel was smoothly narrowed in the lateral direction from 0.250 to 0.165 m; the width of the discharge region was 0.165 m. The distance between the cathode array and the anode was 33 mm.

The cathode array was made of a fiberglass laminate board, on which the knife edge cathode elements were mounted, insulated from one another by a layer of boron alumonitrate. The thickness of the protective layer was 3.5 mm and that of the fiberglass board, 5 mm. In a single row, eighteen nickel knife-blade cathodes with a working surface area of 40×35 mm were mounted along the flow direction in a row, separated from one another by a distance of 20 mm. Three rows of cathodes were mounted on the cathode, separated by a distance of 50 mm between their axes of symmetry transverse to the flow. Figure 1 is a sketch of the discharge chamber with a single longitudinal series. The cathodes were connected through individual ballast resistors to the power supply. The plane anode, which occupied the entire lower wall of the chamber was made of copper. The potentials of and currents to the cathodes were measured on the middle row of cathodes while all three rows were working. The temperature of the cathode array was measured with the aid of thermocouples located at the points where the fiberglass substrate and the base of a cathode element came into contact. The temperature of the gas was measured with a thermocouple on the axis of the channel a distance of 0.3 m past the outlet cross section of the discharge region.

Figure 2 shows a sketch of the discharge chamber used for experiments with a gas feed into the discharge region. Between each pair of transverse rows of cathodes a slit with a width of 0.15 mm was made at a distance of 7 mm behind the first cathode along the flow.

During the experiments the static air pressure in the discharge chamber was measured and the mass feed of air in the channel was monitored. At a distance of 275 mm ahead of the discharge region a structure for mounting a turbulizer was provided. In the experiments a turbulizer was used that consisted of eight cones which uniformly filled the transverse cross section of the channel. The cones had bases with diameters of 25 mm and their height was 37 mm.

SYSTEM OF GAS DYNAMIC EQUATIONS

The system of gas dynamic equations (1)–(5) describes the flow of the gas in the approximation of a narrow channel and includes the steady state Prandtl equations,[6] the equation of state of the gas, and the vibrational kinetic equations (two temperature model). The anharmonicity of the nitrogen molecules, the diffusion of vibrationally exited molecules, and the accommodation of their energy at the channel walls were taken into account. In the calculations we used an implicit six-point scheme with second order accuracy with a rectangular grid.[7] The equations are

$$\frac{\partial \rho u}{\partial x}+\frac{\partial \rho v}{\partial y}=0, \tag{1}$$

$$\rho u\frac{\partial u}{\partial x}+\rho V\frac{\partial u}{\partial y}=-\frac{dp}{dx}+\frac{\partial}{\partial y}\left(\mu_{\text{eff}}\frac{\partial u}{\partial y}\right), \tag{2}$$

$$c_p\rho u\frac{\partial T}{\partial x}+c_p\rho v\frac{\partial T}{\partial y}=jE\alpha+\Theta N\frac{e-e_0}{\tau}+\Theta N Q_{\text{ang},N}+\frac{\partial}{\partial y}\left(\lambda_{\text{eff}}\frac{\partial T}{\partial y}\right), \tag{3}$$

$$p=\rho RT, \tag{4}$$

$$u\frac{\partial e}{\partial x}+v\frac{\partial e}{\partial y}=\frac{jE(1-\alpha)}{\Theta N}-\frac{e-e_0}{\tau}-Q_{\mathrm{ang},N}+\frac{\partial}{\partial y}\left(D_{N,\mathrm{eff}}\frac{\partial e}{\partial y}\right). \quad (5)$$

Here x,y are the axes of a Cartesian coordinate system (Fig. 1); u,v are the components of the velocity vector along the x and y axes; c_p, ρ, p, T, and N are the specific heat, density, pressure, temperature, and molecular density of the gas; μ_{eff} is the effective coefficient of viscosity ($\mu_{\mathrm{eff}}=\mu+\mu_{\mathrm{tur}}$), where μ is the dynamic viscosity and μ_{tur} is the turbulent viscosity; λ_{eff} is the effective thermal conductivity ($\lambda_{\mathrm{eff}}=\lambda+\lambda_{\mathrm{tur}}$), where λ is the coefficient of molecular thermal conductivity and λ_{tur} is the coefficient of turbulent thermal conductivity; e is the average number of vibrational quanta with energy $k\Theta$ per gas molecule and stored in N_2 molecules ($\Theta=3360$ K); e_0 is the equilibrium value; R is the gas constant of the mixture; k is the Boltzmann constant; jE is the volume energy input into the positive column of the discharge; $Q_{\mathrm{ang},N}$ is the dissipated power of the vibrational energy through the anharmonic channel of N_2 molecules, used in the form given in Ref. 8; $D_{N,\mathrm{eff}}$ is the effective diffusion coefficient of vibrationally excited N_2 molecules ($D_{\mathrm{eff}}=D^*+D_{\mathrm{tur}}$), where D_{tur} is the turbulent diffusion coefficient of the molecules; and, D^* and D are the kinematic coefficients of self-diffusion of the vibrationally excited and unexcited molecules, respectively.

Vibrational excitation of the molecules owing to a change in the resonance exchange cross section can cause a change in the self-diffusion coefficient. For example, according to the data of Refs. 9 and 10, for the CO_2 molecule $D^*/D=0.6-0.7$. For the N_2 molecule $D^*/D=0.5-0.3$, and for O_2, $D^*/D=0.5-0.7$, both at $T=1000-1500$ K.[11] In this paper we take $D^*/D=0.5$. For high populations of the vibrational levels of the N_2 molecule, the coefficients of thermal conductivity, viscosity, and diffusion of air should decrease. As an example, we list values calculated according to formulas for gas mixtures.[12] For $D^*/D=0.5$, $T=300$ K, and a change in the average number of vibrational quanta per molecule of air (e) from 0 to 0.7, the diffusion coefficient D^* increased by a factor of 1.37. The thermal conductivity and viscosity decreased by factors of 1.53 and 1.51, respectively.

For calculations of air flows, the vibrational relaxation time for nitrogen on water molecules was computed using the formula $\tau=0.145\times10^{-20}T\exp(-30.6/T)\xi_{H_2O}$, where ξ_{H_2O} is the volume fraction of water molecules in the air. The fraction of the energy of the discharge going into direct heating of the air, α, was determined from experimental data[13] and approximated using the formula $\alpha=0.002P(x)T_0/T(x,y)+0.04$, where $[p]=$Torr. The boundary conditions for the system of Eqs. (1)–(5) were written in the following form: at the channel walls for $y=0,H$, we have $u=0$, $v=0$. The temperature was specified constant at the walls at $T=T_{\mathrm{st}}$. The degree of filling of the vibrational levels of the nitrogen molecule was obtained using the condition $D_{N,\mathrm{eff}}(\partial e_N/\partial y)=-0.5\gamma_N a e_N/(2-\gamma_N)$. Here γ_N is the accommodation coefficient for the vibrational energy at the channel walls, given by $\gamma_N=0.033$, and a is the thermal speed of the molecules. Uniform distributions over the height of the channel were specified for the flow parameters at the channel inlet: $u(0,y)=u_0$; $T(0,y)=T_0$, $p(0)=p_0$, and $e_N(0,y)=e_0(T_0)$. u_0 was calculated from the mass rate of feed of air through the transverse cross section of the channel, G_0, $u_0=G_0/S/\rho_0$, where S is the transverse cross sectional area of the channel and the subscript 0 denotes values at the inlet cross section of the channel.

The turbulent boundary layer was described using an algebraic model in the framework of a two-layer Clauser scheme. In the interior of the boundary layer the turbulent diffusion coefficient was determined from the Prandtl formula with a van Driest damping factor.[14] In the outer region the turbulent viscosity was calculated using the Clauser formula with a factor to account for the intermittence of the flow near the outer layer of the boundary layer.

The electric field strength E in the positive column of the discharge, the current I_i flowing through the ith cathode, the ballast resistance R_i, the power supply voltage U, and the channel height H are related by the equation

$$\int_0^H E\,dy+I_iR_i+U_K+U_A=U, \quad (6)$$

where U_K is the cathode potential fall and U_A is the anode potential fall.

For calculating the cathode potentials, we assume that the condition $E/N=$const holds in the positive column of the discharge.[15] $E/N=$const implies $E/p_*=$const, where $p_*=p(x)T_0/T(x,y)$. (* means that the quantity is reduced to normal or to the initial gas temperature.) Then Eq. (6) could be written in the form

$$\frac{E}{p_{*i}}pH_{*i}+I_iR_i+U_K+U_A=U, \quad (7)$$

where

$$H_{*i}=H\int_0^H T_0/T(x_i,y)\,dy$$

is the channel height reduced to the gas temperature at the channel inlet, x_i is the coordinate of the ith cathode; and, E/P_{*i} is averaged over the channel height in the zone of influence of the ith cathode.

In the gas inflow regime, a one dimensional system of equations obtained by simplifying the system (1)–(5), specifically for $v=0$ and zero gradients of the flow parameters in the y direction, was solved. The gas inflow was modeled by a gradual increase in the mass feed rate of the gas after each transverse row of cathodes. The heated gas was assumed to undergo instantaneous mixing in the discharge region. The conditions at the channel inlet were written in the form $u(0)=u_0$, $T(0)=T_0$, $p(0)=p_0$, and $e_N(0)=e_0(T_0)$. The cathode array contained 20 slits, one in front of the first cathode row, on after the last row, and between every two transverse rows there was one slit. Thus, u_0 was calculated using the formula $u_0=0.05G_V/\rho_0/S$, where G_V is the mass feed rate of air through the cathode array. The gas in the cavity lying between the cathode array and the delivery ap-

erture controlling the mass feed rate of the air inflow was assumed to be $T_0=300$ K. The pressure was measured directly in the experiments through a static aperture on the axis of the anode. The equation for calculating the temperature of the cathode array was

$$c^*\rho^*\frac{\partial T^*}{\partial t}=\lambda^*\left(\frac{\partial^2 T^*}{\partial x^2}+\frac{\partial^2 T^*}{\partial y^2}\right), \tag{8}$$

where c^*, ρ^*, and λ^* are the specific heat, density, and thermal conductivity, respectively, of the material in the cathode array.

The boundary condition for the temperature at the protective surface was determined from the condition that the heat flux at the gas-cathode array boundary should be constant. The continuous heat flux condition was also satisfied at the interfaces of elements of the cathode array with different thermal and physical parameters. At the working surface of the cathode, a heat flux directed from the cathode fall region into the cathode volume was specified.

EXPERIMENTAL DATA AND NUMERICAL ANALYSIS OF THEM

The temperature of the cathode array was measured in a transverse flow of air with ignition of a discharge from one transverse row of cathodes and an initial gas temperature $T_0=293$ K at the cathode array. The measurements shows that the temperature of the cathode array for a given discharge burning time τ is proportional to the current at the cathode element. It depends weakly on the pressure, gas feed rate, and the characteristics of the flow turbulence. For example, for $G_0=20.6\times10^{-3}$ kg/s, $I_K=73$ mA, $\tau=60$ s, and p_0 ranging from 41 to 62 Torr, the maximum heating of the cathode array varied from from 73.5 to 67 K (9.3%), and for $\tau=600$ s from 150 to 137 K (9%). At the same time, gas on the channel axis was heated from 50 to 75 K (50%), independently of the discharge burning time. The effect of changes in the mass feed rate of the gas was more significant. For $p_0=52$ Torr, $I_K=73$ mA, $\tau=60$ s, and G_0 varying from from 20 to 50 kg/s, the maximum heating of the cathode array changed from 100 to 75 K (20%), and for $\tau=600$ s, from 150 to 135 K (10%). When G_0 was increased from 20×10^{-3} to 50×10^{-3} kg/s, the heating of the gas on the axis of the channel decreased from 70 to 30 K (57%). Installation of the turbulizers, which resulted in complete detachment of the boundary layer ahead of the discharge region and generated a large scale turbulence, had no effect on the cathode temperature. Thus, for a constant current, the gas temperature in relative units depended much more strongly on the initial flow parameters than did the temperature of the cathode array. Thus, the cathode array was heated primarily by heat transfer from the cathode fall region into the cathode volume with subsequent redistribution of the heat over the cathode array. This is also indicated by the low degree of asymmetry in the temperature of the cathode array relative to the location of a cathode. Figure 3 shows calculated and experimental temperature distributions of the cathode array for different discharge burning times ($I_K=110$ mA). A temperature maximum is observed at the site of a cathode element. The calculated values were in satisfactory agreement with the experimental data for currents of 70 and 90 mA.

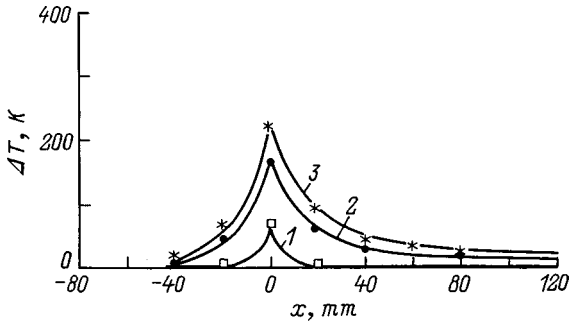

FIG. 3. The distribution of the temperature increase in the cathode array: *, ●, □ — experimental data; smooth curves — calculations; $I=110$ mA; $\tau=20$ (*1*), 180 (*2*), 300 s (*3*).

The temperature distributions were calculated using Eq. (8). The heat flux through the working surface of the cathode was determined in the following approximations. Under the experimental conditions, when the deviation in the current density from the normal value was relatively small, the cathode fall was essentially independent of the current. Estimates show that the most of the heat released in the cathode fall region enters the volume of a cathode element. This is because of the low height of the cathode fall region and is consistent with the experimental results. Thus, as a rough accounting for the energy expended in creating electron-ion pairs the formula for calculating the heat flux can be written in the form $q=\chi\times(U_K-\Omega)\times I_K/S$, where U_K is the cathode potential fall; I_K is the current flowing through a cathode element; S is the area of the working surface of a cathode element; Ω is the cost of an ion; and, χ is the fraction of the heat from the cathode fall region entering the cathode. There is an objective uncertainty in any determination of U_K under particular conditions. This quantity depends, for example, on the state of the cathode surface. Here we did not measure U_K and we took $U_K=208$ V in accordance with Ref. 16. The best agreement between the theoretical and experimental temperatures was obtained for $\Omega=30$ V and $\chi=0.7$ over the entire range of currents, pressures, and flow velocities. Given the approximate character of the previous estimates, we can, in fact, use the simple formula $q=U_\delta\times I_K/S$, where U_δ is the effective cathode potential fall. It is determined from the coincidence of the calculated and experimental data for one arbitrary choice of parameters: the cathode current and the velocity and pressure of the gas. The value calculated in this way can be used over a wide range of initial flow parameters.

This algorithm was used to set the boundary conditions with respect to the temperature on the surface of the cathode array for the calculations of the two dimensional gas flow. Calculations of the distribution of the temperature on the cathode array when a discharge was struck from all the cathode elements showed that it has a sawtooth shape. The temperature maxima occur at the sites of the cathodes. The temperature $T_{y=H}=T(x)$ on the cathode surface was specified for a discharge burning time of 300 s. Its maximum was 600 K. This discharge burning duration was consistent with the time to make the measurements of the currents and po-

TABLE I.

Cathode No.	I_K, mA	U, V	W, kJ/(kg·s)	$G_i\times10^{-3}$, kg/s
1	120	2083	563	1.14
5	131	2408	162	5.1
10	167	2664	116	10.2
17	177	2905	76	18.2

tentials of the cathode elements. The condition $T_{y=0}=T_0$ was specified at the anode.

In the following we present the results of a study of the limiting characteristics of the discharge. The operating characteristics of each cathode were determined by measuring the cathode current at the boundary of the stable discharge regime. The measurements were made in a transverse flow with discharge initiation from one transverse row. An analysis of these data showed that except for two cathodes with reduced currents, the spread in the critical current relative to the average was $\pm10\%$. The deviation in the cathode current for the two exceptions was 30%. For the subsequent studies of the individual characteristics, four cathodes with similar critical currents, Nos. 1, 5, 10, and 17 along the flow, were chosen. In the gas inflow regime, the limiting current and discharge voltage were measured through the cathode array. Measurements were made with discharge initiation from one transverse row with $p_0=48$ Torr, $T_0=293$ K, and $G_0=20.6\times10^{-3}$ kg/s.

Table I lists the limiting currents, voltages, and energy inputs. The specific energy input were calculated using the formula $W=(U-U_K-U_A)I/G_i$, where I is the discharge current and G_i is the mass feed rate of the gas through the transverse cross section of the channel in the zone of influence of the ith cathode element. It is evident from the table that on going from the first to the second cathode, the critical current changed little, but because of the increased gas feed rate the limiting specific energy input decreased by a factor of 3.47. The limiting current for cathode No. 17 was 47.5% higher, but the specific energy input was 7.4 times smaller than for the first cathode.

Thus, at the stability boundary for the discharge, the parameters in the cathode region have a relatively weak dependence on the parameters of the large scale gas flow. This suggests that the instability develops when a critical current is reached in the cathode region. The relatively weak dependence of the cathode region on the flow parameters is a consequence of its small dimensions and the anomalously high electric field strength. The role of volume and surface processes in the development of the instability of the cathode layer has been studied elsewhere.[17–19]

We now proceed to the development of an algorithm for calculating the distribution of the current over the cathode elements. To study the character of the distribution of $E/p_{*i}=\text{const}$ along the discharge vessel we calculated

$$\frac{E}{p_{*i}}=\frac{U_i-U_K-U_A}{pH_{*i}}, \tag{9}$$

where U_i is the experimental value of the potential of the ith cathode.

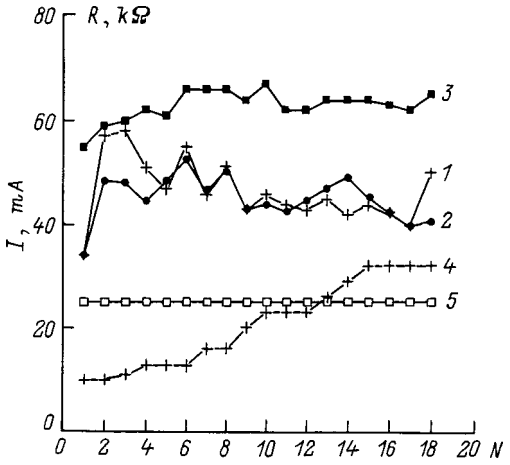

FIG. 4. Distributions of the discharge current over the cathodes and of the ballast resistances: *1–3* — I_i; *4,5* — R_i; *1,2,4* — transverse discharge; *1,3* — experiment; *2* — calculated; *3,5* — discharge with gas blown in through the cathode array; N is the number of the cathode.

In the calculations for an air flow, we took $U_A=100$ V.[20] The calculations were done for a transverse discharge and for operation with a gas inflow through the cathode array. In the calculations of the two dimensional transverse flow in a plane intersecting the electrodes, the lateral constriction of the channel was taken into account by a corresponding increase in the mass feed rate of the gas through unit cross sectional area of the channel.

Figures 3–5 show the distributions of the gas and discharge parameters, the experimentally measured potential difference (U_1-U_i) between the first and ith cathodes, the cathode currents I_i, the ballast resistances R_i, the calculated values of E/p_*, H_*, and I_i, and the temperature rise $\Delta T_i=T_i-T_0$ on the channel axis in the discharge region in the direction of the flow. The calculations and experiments with the transverse discharge and with the discharge with a gas inflow were done under the following conditions: $p_0=58$ Torr, $T_0=293$ K, and $G_0=20.6\times10^{-3}$ kg/s.

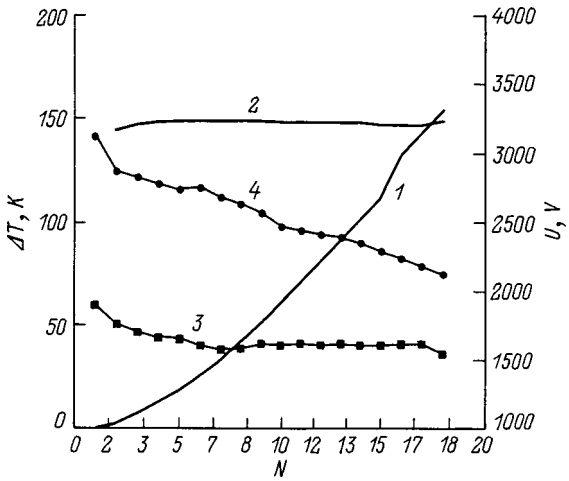

FIG. 5. Distributions of the gas temperature increment in the core of the flow and of the cathode potentials: *1,2* — ΔT_i; *3,4* — U_i; *1,4* — transverse discharge; *2,3* — discharge with gas blown in through the cathode array; N is the number of the cathode.

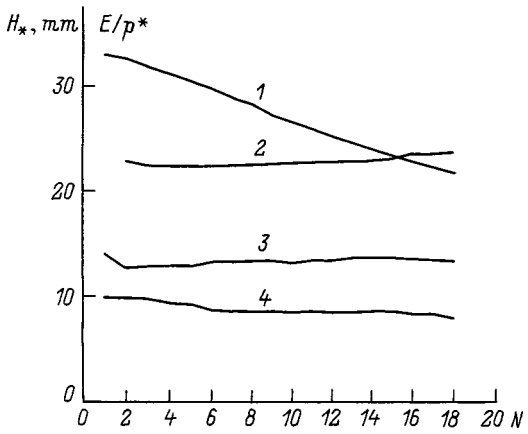

FIG. 6. Distributions over the cathodes of the channel height reduced to the normal gas temperature and of the ratio E/p_*: *1,2* — H_*; *3,4* — E/p_*: *1,3* — transverse discharge; *2,4* — discharge with gas blown in through the cathode array; N is the number of the cathode.

The way in which the ballast resistances must be distributed over the cathodes in order to obtain a uniform distribution of the discharge current over them depends on how the gas flow is organized (Fig. 3). If the gas is blown in, the uniform entry of cold gas into the discharge region levels out the gas temperature in the direction of the flow, while in the transverse discharge regime the temperature increases monotonically in the direction of the flow. Therefore, in the transverse discharge the cathode potentials fall off significantly in the direction of the flow, while in the discharge with gas blown in, they vary relatively little (Fig. 4). Thus, in the first case the ballast resistances were specified to increase downstream and in the second, to be the same for all the cathodes. Calculations showed that in the transverse flow the increment in the gas temperature on the channel axis in the zone of influence of the last cathode was 154 K, while the height of the channel, reduced to the temperature at the channel inlet, decreased substantially, from 33 to 21.7 mm. The values of E/p_{*i} corresponding to the different cathodes differed from the average $E/p_* = 13.2\ \mathrm{V/(cm \cdot Torr)}$ obtained by averaging over cathodes No. 2–17 by only 3% (Fig. 5). This value of E/p_* is consistent with data for a transverse discharge in atmospheric air.[1] The first and last cathodes were not included in averaging E/p_{*i} since they are somewhat anomalous.

In the regime with gas blown in, the spread in the calculated values of E/p_{*i} over the cathodes is somewhat greater, probably because of the inadequate one dimensional model for the complicated vortical flow in the channel. When gas was blown in through the cathode array, the difference in E/p_{*i} from the average $E/p_* = 8.81\ \mathrm{V/(cm \cdot Torr)}$ reached 10% only for the second cathode, and was less than 6% beginning with the third. Note that with gas blown in, the mass feed rate of the gas through the transverse cross section of the channel increased from the first to the last cathode by roughly a factor of 20. Against this background the relatively weak variation in E/p_{*i} justifies the use of the approximation $E/p_{*i} = \mathrm{const}$ for the positive column. However, this parameter depended on the way the discharge is organized, which may explain, for example, the effect of the electrode regions. When the length L of the Faraday dark space is taken into account, Eq. (9) can be rewritten as

$$\frac{E}{p_{*i}} = \frac{(U_i - U_K - U_A)}{\left(p(H-L)\int_0^{H-L} T_0/T(x_i, y)\,dy\right)}. \tag{10}$$

When gas is blown in through the cathode array, gas is removed from the Faraday dark space toward the anode. By analogy with a longitudinal discharge, it may be assumed that this causes the Faraday dark space to become longer.[21] Equation (10) implies that we obtain a low value of E/p_{*i} if the elongation of L is not taken into account.

Figure 6 shows the calculated cathode currents. The current at the first cathode was specified in the calculations, the power supply voltage was determined from Eq. (6) using Eqs. (1)–(7) for an average $E/p_* = 13.2\ \mathrm{V/(cm \cdot Torr)}$, and the distribution of the current over the cathodes was computed. Evidently, the calculated and experimental currents are in satisfactory agreement for most of the cathodes.

The potentials and currents calculated using the proposed algorithm can be used to set the boundary conditions for a more detailed calculation of the discharge. For example, two dimensional calculations of the positive column of a discharge have been done[22] with one dimensional gas dynamics and without a consistent calculation of the cathode layer. The lack of a consistent calculation of the cathode layer means that additional conditions have to be imposed on the working surface of the cathode. A complete and consistent calculation of the discharge, especially for a large number of cathodes, is, so far, a separate problem.

CONCLUSION

We have studied the parameters of discharges fundamentally different gas dynamic structures in transverse flows and in flows with a distinct vortical structure. A numerical analysis of the experimental data showed that for both methods of organizing the flow in the positive column of the discharge, E/N (averaged over the channel height) depends weakly on the distance to the inlet cross section of the channel. The distribution of the ballast resistances for producing a uniform current distribution over the cathode elements depends on the way the flow is organized. This behavior permits a fairly simple calculation of the flow parameters and the distribution of the current and discharge voltage over the cathode elements.

The above results imply that the state of the cathode region determines the thermal regime of the cathode array and the stability of the discharge. It has been shown that the distribution of the current over the cathodes depends on the flow parameters. At the same time, the temperature profiles of the gas flow depend on the thickness of the boundary layer, the turbulence characteristics, the electrode temperatures, and the channel height and geometry. Thus, in some cases, the ambiguous influence of turbulence on the discharge stability might be explained if these interrelations were consistently taken into account. The results obtained in this paper can serve as a model for calculating the limiting parameters of the discharge.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 95-01-00619).

[1] E. P. Velikhov, V. S. Golubev, and S. V. Pashkin, Usp. Fiz. Nauk **137**, 117 (1982) [Sov. Phys. Usp. **25**, 340 (1982)].
[2] A. V. Bondarenko, V. S. Golubev, E. V. Dan'shchikov, Fiz. Plazmy **5**, 687 (1979) [Sov. J. Plasma Phys. **5**, 386 (1979)].
[3] A. I. Ivlyutin and Yu. S. Levitan, Fiz. Plazmy **17**, 720 (1991) [Sov. J. Plasma Phys. **17**, 422 (1991)].
[4] E. Armandillo and A. S. Kaye, J. Phys. D **13**, 321 (1980).
[5] A. V. Astakhov, G. A. Baranov, Yu. B. Butaev *et al.*, NIIÉFA Preprint No. A-0494 [in Russian], D. V. Efremov Scientific-Research Institute of Electrophysical Apparatus, Leningrad (1980), 24 pp.
[6] H. Schlichting, *Boundary Layer Theory*, 6th ed. [McGraw-Hill, New York (1968); Nauka, Moscow (1974), p. 711].
[7] V. M. Paskonov, V. I. Pozhelaev, and L. A. Chudov, *Numerical Modeling of Heat and Mass Transfer Processes* [in Russian], Nauka, Moscow (1984), 284 pp.
[8] B. F. Gordiets, A. I. Osipov, an L. A. Shelepin, *Kinetic Processes in Gases and Molecular Lasers* [in Russian], Nauka, Moscow (1980), 512 pp.
[9] M. Kovaes *et al.*, J. Chem. Phys. **50**, 411 (1969).
[10] X. Doyenette, J. Chem. Phys. **60**, 697 (1974).
[11] W. F. Ahtye, J. Chem. Phys. **57**, 5542 (1972).
[12] K. M. Aref'ev, *Transport Phenomena in Gases and Plasmas* [in Russian], Énergoatomizdat, Moscow (1983).
[13] Yu. S. Akishev, S. V. Dvurechenskiĭ, A. I. Zakharchenko *et al.*, Fiz. Plazmy **7**, 1273 (1981).
[14] Yu. V. Lapin, *Turbulent Boundary Layers in Supersonic Gas Flows* [in Russian] (1982), 312 pp.
[15] W. L. Nighan, in *Principles of Laser Plasmas*, edited by G. Bekefi [Wiley, New York (1973), p. 272; Énergoatomizdat, Moscow (1982), p. 269].
[16] Yu. P. Raizer, *Physics of Gas Discharges* [in Russian], Nauka, Moscow (1987), 592 pp.
[17] Yu. D. Korolev and G. A. Mesyats, *Physics of Pulsed Breakdown in Gases* [in Russian], Nauka, Moscow (1991), 224 pp.
[18] Yu. S. Akishev, A. P. Napartovich, V. V. Ponomarenko *et al.*, IAÉ, Preprint No. 3900/7 (1984), 24 pp.
[19] Yu. S. Akishev, K. V. Baiadze, V. M. Vetsko *et al.*, Fiz. Plazmy **11**, 999 (1985) [Sov. J. Plasma Phys. **11**, 582 (1985)].
[20] Yu. M. Belyakov, R. I. Galeev, F. M. Gaĭsin *et al.*, Teplofiz. Vys. Temp. **17**, 1173 (1979).
[21] F. I. Vysikaĭlo, Teplofiz. Vys. Temp. **24**, 657 (1986).
[22] R. Sh. Basyrov, F. M. Gaĭsin, A. M. Minnigulov, and B. A. Timerkaev, Teplofiz. Vys. Temp. **32**, 334 (1994).

Translated by D. H. McNeill

Effect of the gas dynamic structure of a flow on the parameters of a self-sustained discharge. II

G. A. Baranov and S. A. Smirnov

D. V. Efremov Scientific-Research Institute of Electrophysical Apparatus, 189631 St. Petersburg, Russia
(Submitted October 15, 1998)
Zh. Tekh. Fiz. **69**, 49–55 (November 1999)

The limiting current of a self-sustained glow discharge is calculated. Two-dimensional equations for the flow of a viscous, vibrationally nonequilibrium gas and a model of the cathode sheath are used. The validity of the approximations which form the basis of the cathode sheath model was tested with experimental data for anomalous and normal currents. The effects of laminar and turbulent gas flow and of the geometric dimensions of the channel on the limiting discharge current are examined. © *1999 American Institute of Physics.* [S1063-7842(99)00911-3]

INTRODUCTION

Instability of a transverse discharge with a sectioned cathode array can show up in the form of brightly luminous trails growing from some of the cathodes. The criticality of conditions in the cathode region for the development of a discharge instability has been demonstrated repeatedly. Contractions, for example, develop as a result of field emission,[1] nonuniformities in the oxide layer on the cathode,[2] and increases in the collisional ionization and secondary emission coefficients resulting from an increase in the density of metastable particles.[2]

In order to develop a physical and mathematical model for the effect of the gas dynamic structure of the flow on the discharge parameters, it is appropriate to conduct experiments and analyze them numerically under conditions such that the initial gas temperature changes at the inlet to the discharge chamber, while the density and velocity of the gas are fixed. Then the initial gas temperature determines the energy input to the discharge necessary to reduce the gas density by a certain amount. Thus, varying the initial gas temperature makes it possible to vary the degree of rarefaction of the gas for a fixed energy input and constant initial conditions throughout the entire temperature range. Against this background, it is easier to compare the structure of the gas flow with the power input level and discharge regime. In this paper we have not set ourselves the task of making a detailed analysis of the mechanism for the discharge instability. As criterion for the development of the discharge instability we have taken the attainment within the cathode dark space of a maximum value of E/N. Note that for a nonuniform gas temperature the maximum in E/N can lie at some distance from the cathode surface. The distribution of the discharge current over the cathode elements was calculated using the algorithm developed in the preceding article.[3] This model was used to analyze experimental data from Ref. 4.

To a certain extent, the condition $(E/N)_{\max}=\text{const}$ is analogous to the condition $E_{y=0}=\text{const}$ used for identifying the instability.[1,5] In Ref. 5, a discharge on a palladium cathode made of a wire with a diameter of 0.74 mm enclosed in an aluminum oxide tube. $E_{y=0}$ was not measured there[5] but its value was determined by estimating the magnitude of the potential fall in the cathode dark space and its thickness. The gas pressure was normalized with respect to the temperature averaged over the height of the cathode dark space, since the parameters of the latter depend on the gas density, rather than on the absolute magnitude of the pressure. Starting with $E_{y=0}=\text{const}$, which is satisfied for the critical currents at pressures $P_0=25-150$ Torr, we concluded that a field emission instability develops.

More exact calculations show that normalizing the gas pressure with respect to the average temperature is a crude approximation for determining the field at the cathode surface. The temperature of the cathode surface varies little compared to the remaining part of the cathode dark space as the current is varied. Thus, the field at the cathode for a given pressure was found[5] to be essentially independent of the temperature distribution of the gas in the cathode dark space. It might be determined by the gas pressure or the degree of anomaly in the current. We denote the magnitude of the field, determined by analogy with Ref. 5, by E_S. The thickness of the cathode dark space, d_S, which is defined in terms of the average density N_S, satisfies the condition $N_S d_S \approx \text{const}$. Given the linearity of the field distribution in the cathode dark space, we have $E_S \sim 2U_K/d_S = 2U_K N_S/(N_S d_S) \approx \text{const}\, N_S$, where $U_K \approx \text{const}$ is the cathode fall. Thus, in Ref. 5, the condition $E_S \approx \text{const}$ should imply $E_S/N_S \approx \text{const}$, where E_S is a scale for the field in the cathode dark region.

MODEL FOR THE CATHODE FALL REGION WITH A NONUNIFORM GAS TEMPERATURE

The preceding discussion shows the appropriateness of developing a model for a cathode sheath with a nonuniform gas temperature. Modelling the cathode region and making qualitative estimates of its parameters are done with different degrees of complexity and consistency in the approximations. Here we obtain equations for estimating the parameters

in the cathode fall region using the classical similarity relations, modified to account for a nonuniform gas temperature. The degree of reliability of the results is tested by comparing the computational results with experiment.

Let us briefly discuss the main methods for modelling the cathode sheath. The nonlocality of the electron energy distribution function owing to the strong nonuniformity of E/N in the cathode layer requires that a kinetic equation be solved for the electrons. The most complete numerical study of the electron distribution function in the cathode dark space can be done using a Monte Carlo method. It has been used[6] for calculating the electron distribution function for high fields. For low fields a diffusion- drift approximation is used. Calculations have been done for the cathode dark space and the negative glow. Monte Carlo methods are rather complicated and time consuming. The most accessible method for solving the kinetic equation has been proposed in Ref. 7. This method also describes nonlocalization effects, but it can also be used to model the hydrodynamic approximation and electron diffusion. It has been used[8] to calculate the layered structure of a cathode sheath consisting of a cathode dark space, a negative glow, and a faraday dark space. The changes in the field on going from the cathode dark space to the positive column were shown to be nonmonotonic. Two groups of electrons were obtained in the cathode dark space and the negative glow: fast and slow. As in Ref. 6, because the atoms are ionized by a beam of fast electrons, the maximum ionization rate lies outside the cathode dark space.

In Ref. 9 the nonlocalization of the electron distribution function was taken into account by renormalizing the local Townsend coefficient,

$$\frac{\alpha}{p}=Ae^{(-Bp/E)}. \tag{1}$$

In order to account for the dependence of the ionization rate on the field over the interval $x_I=(x-\Delta x_I,x)$, where

$$e\int_{x_I}^{x}E\,dx'=I \tag{2}$$

is the ionization potential, the ionization rate was written in the form

$$Q(x)=\mu_e(n_e|E|)_{x=x-\Delta x_I}e^{-Bp/|E(x)|}. \tag{3}$$

The equations for the field strength and current density are written in the form

$$\frac{dE}{dx}=4\pi e(n_e-n_i), \tag{4}$$

and

$$\frac{dj_e}{dx}=\frac{dj_i}{dx}=Q-\beta n_e n_i, \tag{5}$$

where E is the electric field strength, n_e, n_i, j_e, and j_i are the densities and current densities of the electrons and ions, and β is the recombination coefficient.

For current densities close to normal, the calculations showed that electron diffusion has a significant effect on the current voltage characteristics, while the effect of nonlocalization was substantially less.

In Ref. 10 the effect of nonlocalization has been taken into account by introducing a time delay between an electron's gaining the energy required for ionization and the moment of ionization. This was done by a suitable renormalization of the Townsend coefficient and changing the form of the original equation,

$$\frac{dj_e}{dx}=\alpha j_e(x). \tag{6}$$

Because the nonlocalization beyond the confines of the cathode dark space was taken into account, the calculations of Refs. 9 and 10 yielded a local minimum in the electric field strength. For nitrogen, they showed[10] that agreement between the potential fall U_N and the experimental data can be obtained by choosing the secondary emission coefficient γ. This was not possible for the normal current density j_N. Note that the problem of choosing γ arose in all these papers,[6–10] but the problem of calculating the normal current density in a self-consistent manner has not yet been definitively solved.

Therefore, by renormalizing the Townsend coefficient in the diffusion- drift [or, simply, drift (5)] approximation it is possible to determine several effects associated with the nonlocalization of the electron distribution function. At the same time, the diffusion-drift approximation with a local Townsend coefficient yields fair agreement with the experimental current-voltage characteristic and field distribution in the cathode dark space.[11,12] In principle, this corresponds to the classical Engel-Steenbeck theory,[13] which is still widely used for estimating the parameters of the cathode sheath: the cathode potential fall, the current density, and the thickness of the cathode dark space. The major theses of the Engel-Steenbeck theory have not been refuted by the results of more complex, modern techniques. The latter make it possible to compute the fine structure of the cathode sheath, but they are still complicated and do not guarantee a satisfactory result.

Let us introduce the main propositions of the Engel–Steenbeck theory in more detail as we require them in order to develop the model of the cathode dark space used here. According to that theory, Eq. (5) is satisfied for a one-dimensional cathode layer with a uniform gas density in the steady state, along with the conditions

$$j_e+j_i=j;\quad j_{eK}=\gamma j_{iK}=\left[\frac{\gamma}{1+\gamma}\right]j, \tag{7}$$

where j is the current density in the cathode dark space and j_{eK} and j_{iK} are the electron and ion current densities at the cathode.

Equations (6) and (7) yield a condition for self-sustainment of the discharge,

$$\int_0^d\alpha[E(x)]dx=\ln\left(1+\frac{1}{\gamma}\right). \tag{8}$$

In the theory, one solves the system of Eqs. (1), (4), and (6)–(8) for $n_i \gg n_e$ and for a linear distribution of the field in the cathode dark space,

$$E(x) = E_K\left(1 - \frac{x}{d}\right). \tag{9}$$

The distribution (9) has been confirmed by probe measurements, as well as by detailed calculations in some of the papers cited above. The solution of Eqs. (1) and (6)–(9) is not given in elementary functions and in order to obtain an intuitive result, the approximation

$$E(x) = \text{const}, \tag{10}$$

is used instead of Eq. (9) for the cathode dark space. Using this expression yields more or less the same parameters for the cathode layer[14]

$$\frac{E_K}{p} = \frac{B}{C + \ln pd}, \quad U_K = \frac{Bpd}{C + \ln pd},$$

$$C = \ln \frac{A}{\ln(1 + 1/\gamma)}. \tag{11}$$

The similarity relations

$$U_N = \text{const}, \quad \frac{j_N}{p^2} = \text{const}, \quad d_N p = \text{const}, \tag{12}$$

hold for the normal current density, where d is the thickness of the cathode dark space, U is the cathode potential fall, and the subscript N corresponds to values for the normal current density.

When the current density deviates from normal, the parameters of the cathode dark space can be calculated using formulas; for the dimensionless quantities

$$\bar{U} = U/U_N, \quad \bar{E} = (E/p)/(E_N/p),$$

$$\bar{d} = (pd)/(pd)_N, \quad \bar{j} = j/j_N \tag{13}$$

Eqs. (11) can be written in the form

$$\bar{j} = \frac{1}{\bar{d}(1 + \ln \bar{d})^2}, \quad \bar{U} = \frac{\bar{d}}{1 + \ln \bar{d}}, \quad \bar{E} = \frac{1}{1 + \ln \bar{d}}. \tag{14}$$

We shall be using Eqs. (9) and (12)–(14) in developing a model of the cathode dark space for a nonuniform gas density. In all the previously cited papers, the gas density was assumed constant in the cathode sheath. This is true for low gas pressures. Our estimates for a nitrogen-copper cathode system showed that neglecting the heating of the gas in the cathode dark space leads to errors in the determination of the current density at the cathode. For example, at pressures of 1, 5, 10, and 37.5 Torr the ratio j/j_N equals 0.864, 0.761, 0.49, and 0.271, respectively, where j is the true current density at the cathode spot and j_N is given by Eq. (12). Heating of the gas affects the cathode fall even in the anomalous regime.

There are few papers devoted to the study of the cathode sheath with a nonuniform gas density. The effect of the gas temperature on the state of the cathode sheath has been studied, for example, in Refs. 15–19. An analytic formula for calculating the normal current density was derived in Ref. 15. To derive it, a similarity criterion from Eq. (12) was used and the gas density was reduced to the temperature averaged over the height of the cathode dark space. The thermal conductivity of the gas was specified to have a temperature dependence of the form $\lambda \propto T^{0.5}$. A nonstationary cathode sheath was studied in Ref. 16 using a continuity equation for the current and the Poisson equation for the electric field. The Euler gas dynamic equations were solved in a simplified form. A matched system of equations for the continuity of the current and the motion of a viscous thermally conducting gas, as well as an equation for the distribution of the electric field, were solved in Eq. (17).

Our derivation of the equation for the field distribution in the cathode dark space and its other parameters is based on an analysis of quantities at slices of the cathode dark space of length $\delta y_{i+1/2} = y_{i+1} - y_i$, bounded by fictitious cross sections with numbers $i = 1, \ldots, M$, where y is the longitudinal coordinate in the cathode region. We shall assume that, as the gas is heated, the cathode dark space transforms and its fictitious cross sections move. In the normal regime, they are ''bound'' to a value of the potential and in the anomalous regime, to the same value of the potential, but changed by the integrated anomaly coefficient. For each segment of the cathode dark space with the normal current density the coordinate and field undergo the transformations

$$\frac{E^N_{i+1/2}}{N^N_{i+1/2}} = \text{const}, \quad \delta y^N_{i+1/2} N^N_{i+1/2} = \text{const}. \tag{15}$$

These transformations conserve the form of Eq. (6), i.e.,

$$dj/dy = \alpha j = jN(y)A \exp[-BN(y)/E(y)].$$

In fact, with Eq. (15), we obtain

$$dj/dy = j dy^*/dy \cdot N^* A \exp[-BN^*/E^*]$$

or

$$dj/dy^* = jAN^* \exp[-BN^*/E^*] = \alpha^* j.$$

The scale for the normal current density[14] was obtained from the approximations $j = en_+\mu_+E$ and $n_+ \approx (4\pi e)^{-1} dE/dy$. We shall write the local derivative of the field in the form $dE/dy \sim \delta\varphi/(\delta y)^2$. For a small degree of anomaly, the potential difference between two cross sections obeys $\delta\varphi = \text{const}$. Using Eq. (15) for the local scale of the normal current density gives $j_N = \mu_+ E^* N(y)^3 \delta\varphi/(dy^*)^2 \propto N^2(y)$. Thus, for a nonuniform gas density, the transformations (15) make it possible to retain the form of all most all the initial equations from the Engel-Steenbeck theory used to derive Eqs. (11), (12), and (14). In addition, Eq. (15) implies that

$$\int_{y_1}^{y} E dy = \int_{y_1^*}^{y^*} NE^*/N^* dy^* (N^*/N) = \int_{y_1^*}^{y^*} E^* dy^* = I,$$

i.e., as the length of the segment between two fixed cross sections of the cathode dark space changes, the energy gained over that segment by the electrons does not vary. Therefore, a similarity condition holds for Eq. (2). The initial

approximations are also consistent with Ref. 18, in that the electron distribution function depends on the potential and local field.

The following propositions have also been formulated for developing an algorithm for calculating the parameters of cathode dark spaces with a nonuniform gas density and different degrees of local anomaly.

1. If the current at a cathode occupies and area than the working surface of the cathode, then the principle of the minimum cathode potential fall is satisfied, i.e., the cathode spot occupies an area such that the cathode fall is minimal.

2. The local value of the normal current density obeys the condition $j^N_{i+1/2} \propto N^2$.

3. Local analogs of the parametric equations (14) are satisfied. They can be used to find the corrections $\overline{E}$ to the local field strength when the local current density deviates from the local normal value; i.e., $\overline{j} = j_{i+1/2}/j^N_{i+1/2}$, $\overline{j} = 1/\overline{\delta}/(1+\ln\overline{\delta})^2$, and $\overline{E} = 1/(1+\ln\overline{\delta})$. Here $j_{i+1/2}$ is the actual value of the current density.

For calculating the parameters of the cathode dark space using our model, it is first necessary to calculate the local values $\overline{j}_{i+1/2}$; they are used to calculate the local corrections for the anomaly, $\overline{\delta}_{i+1/2} = \overline{\delta}(\overline{j}_{i+1/2})$ and $\overline{E}_{i+1/2} = \overline{E}(\overline{\delta}_{i+1/2})$. The normalized segments of the cathode dark space resulting from the deviation from the normal value are defined as $\delta y_{i+1/2} = \delta y^0_{i+1/2}\overline{\delta} N_0/N$, where the subscript 0 corresponds to parameters in the initial uniform gas density in the cathode dark space and to its normal parameters. Now the cathode sheath thickness d can be calculated from the equation

$$\sum_{i=1}^{M-1} \delta y^0_{i+1/2} = d_N = \sum_{i=1}^{M-1} \delta y_{i+1/2} N/(\overline{\delta} N_0)$$

or, in the limit $M \to \infty$,

$$\delta y_{i+1/2} \to 0 \int_0^d N(y)/[N_0 \overline{\delta}(y)]dy = d_N .$$

When the gas density deviates from a uniform distribution, the cross section with coordinate y_0 shifts into position y, where the relationship between these coordinates is determined by

$$\int_0^y N(y)/[N_0 \overline{\delta}(y)]dy = y_0 .$$

Similarly, the local electric field strength is determined from the formula $E_{i+1/2} = E^0_{i+1/2}\overline{E} N/N_0$, while the distribution of the field in the cathode dark space, given Eq. (9) and the normalized coordinates of the fictitious cross sections, is given by

$$E(y) = \frac{2U_N}{d_N}\frac{N(y)}{N_0}\overline{E}(y)\left(1 - \int_0^y \frac{N(y)}{N_0 \overline{\delta}(y) d_N} dy\right). \qquad (16)$$

In the steady state case, the density, pressure, and temperature are related by the equation of state $p = NkT$. The temperature distribution is calculated using the equation

$$\frac{d}{dy}\left(\lambda(T)\frac{dT}{dy}\right) + Ej = 0. \qquad (17)$$

Outside the cathode dark space it was specified that $E/p = 0.1\,\mathrm{V/(cm \cdot Torr)}$ and the pressure in the cathode sheath was uniform over its height. The boundary condition at the cathode surface was specified in the form $T_{y=0} = T_i$, and T_K is the temperature of the working surface of the cathode. At the upper boundary of the computational region, a boundary condition was specified from calculations of the temperature in the boundary layer of the flow with the aid of the system of equations from Ref. 3.

Satisfactory agreement has been obtained[19,20] between calculated normal current densities for a copper cathode and nitrogen using Eqs. (16) and (17) and experimental values.[21,22] The calculations were done for pressures of 37.5-300 Torr and yielded values in agreement with the data of Ref. 23, to within the experimental error, for air and a copper cathode at $p = 22.5-45$ Torr. For an $N_2 + He$ mixture and a copper cathode, good agreement with experiment[23] was obtained for $p = 79.5$ Torr. The calculations were done with the mobility of the nitrogen ions essentially depending strongly only on the nitrogen density. This made it possible to replace the total density of atoms and molecules in the mixture with the density of nitrogen molecules. Note that the form of the distribution of the ratio $j/j_N(y)$ in the cathode dark space, where j is the current density at the cathode spot and $j_N(y)$ is the local value of the normal current density, implies that for $T_{y=0} \leqslant T_\infty$ near the cathode, $j/j_N(y) < 1$, while near the outer boundary of the cathode dark space, $j/j_N(y) > 1$, i. e., for $p > 1$ Torr one can speak of a normal current density only conditionally. A subnormal regime exists near the cathode and an anomalous regime, near the outer boundary of the cathode dark space. The distribution of these zones is such that the principle of a minimum cathode fall is satisfied in the sum.

Calculations of the cathode fall in highly anomalous discharges are also given in Refs. 19 and 20. The calculated values are in satisfactory agreement with experiment.[24] The experiments were done at $p = 23$, 40, and 90 Torr. The maximum cathode fall for nitrogen at $p = 23$ Torr and a copper cathode was about 1500 V. This shows that the model can be used for estimating the parameters of a cathode sheath with a nonuniform gas temperature over a wide range of pressures. The initial values of U_N, j_N/p^2, and pd_N in the calculations were chosen in accordance with experimental data from Ref. 14. There are no published data on pd_N for nitrogen and a copper cathode, so it was chosen from coincidence of the calculated and experimental normal current densities. The choice was made for an arbitrary pressure.

CALCULATING THE LIMITING DISCHARGE CURRENT AS A FUNCTION OF THE GAS TEMPERATURE

The effect of the temperature of longitudinal and transverse flows of dry air and commercial nitrogen ($\sim 2\% O_2$) on discharge uniformity has been studied.[4] In the transverse discharge the anode was made in the form of a copper plate and the ten copper cathodes with a working area of $0.040 \times 0.15\,\mathrm{mm}^2$ were mounted 30 mm apart along the flow and connected to a power supply through a resistance of $18\,\mathrm{k\Omega}$. The discharge gap was chosen to be 30 mm. The gas tem-

perature at the inlet to the channel was varied with the density and velocity of the gas flow held constant. The discharges in a flow of commercial nitrogen with an velocity of 90 m/s, averaged over the cross section, and a pressure of 80 Torr, normalized to standard temperature, were nonuniform for gas temperatures below 190−200 K. At temperatures of 200−240 K a uniform discharge developed and a further increase in the temperature led to a rise in the transition current from a uniform to a nonuniform discharge from 0 to 1.3 A. In experiments with a change in temperature from 160 to 350 K the discharge voltage changed imperceptibly. This is explained by the fact that in the positive column of the discharge, the condition $E/N=\text{const}$ holds, while the gas density was kept constant at the channel inlet.

The experiments of Ref. 3 showed that the discharge instability was initiated in the cathode sheath. Given the similarity of the discharge chambers in Refs. 3 and 4, in analyzing the experimental data of Ref. 4 we assume that the discharge instability developed in the cathode fall regions. In the algorithm for calculating the limiting discharge current, a gradual increase in the number of functioning cathodes was taken into account, consistent with the increase in the total discharge current. By analogy with the real process, it was assumed that after breakdown, the discharge ignites initially at the last cathodes. Then, as the power supply voltage increases, the currents at the functioning cathodes increase, as does the voltage on the idle cathodes that are located upstream. After the potential on the idle cathodes reaches the breakdown value, the gas breaks down from the auxiliary upper cathode. Simultaneously with the new breakdown, the power supply voltage drops, since there is increased heating of the gas in the discharge regions located downstream, so that the electric field strength decreases. This process of successive upstream ignition of the discharge continues until the critical conditions for development of a discharge instability are realized at one of the cathodes.

It was assumed that an instability of the cathode fall region develops after a certain value of E/N, which is the same for all initial gas temperatures, is attained there. The rates of the volume processes involving the electrons are dependent on E/N; under typical conditions they do not depend on the electron density and, therefore, on the current density.[25] The critical value of E/N was assumed to be independent of the gas temperature because the gas density and, therefore, the initial conditions in the cathode region, did not depend on the initial temperature.

The distributions of the temperature of nitrogen in the discharge region and of the currents over the cathodes were calculated using the equations of nonequilibrium gas dynamics for $E/p^*=28.8$ V/(cm·Torr). For this value of E/p^*, the discharge voltage lay within the limits of error of the experimental data.[4] By analogy with our experiments in air, the breakdown field was specified to be 1.32 times the working value. The parameters of the cathode fall region and the distribution of the gas temperature within this region were calculated using Eqs. (16) and (17). As in Ref. 19, it was assumed that the cathode dark space expands in the direction of the anode, in this case at an angle of 45°.

Figure 1 shows the distributions of the gas temperature in the cathode region for two values of the initial nitrogen temperature, $T_0=150$ and 350 K, for a current $I=250$ mA and a pressure, reduced to standard temperature, of $p=80$ Torr. Evidently, the lower the initial temperature is, the greater the overheating of the gas in the cathode layer, since, as the initial temperature is lowered, a smaller temperature change is needed to produce a change in the gas density by a specified amount, while the more the gas is neutralized in the cathode sheath, the higher the degree of its anomaly and the higher the level of Joule heating will be. Figure 2 shows the distribution of the ratio of the electric field strength in the cathode fall region to the local pressure reduced to standard temperature, $(E/p^*)_K$. The parameters are E/N and $(E/p^*)_K$ uniquely related to one another. As opposed to the linear distribution of the field in the cathode dark space observed for a constant gas density, when the gas density is nonuniform $(E/p^*)_K$ develops a maximum inside the cathode fall region, rather than at the cathode surface.

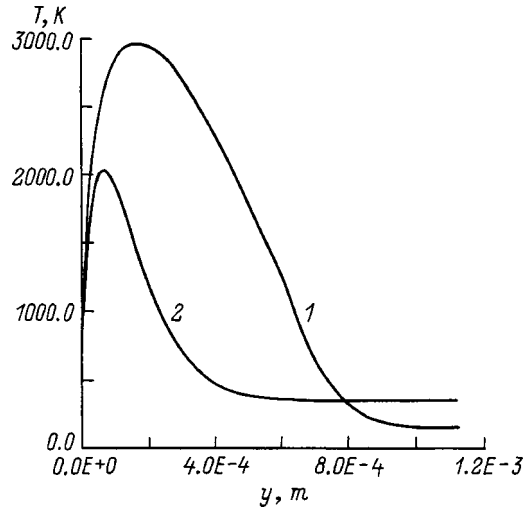

FIG. 1. The distribution of the gas temperature in the cathode fall region: $p^*=80$ Torr, $I=0.23$ A; $T_0=150$ (*1*), 350 K (*2*).

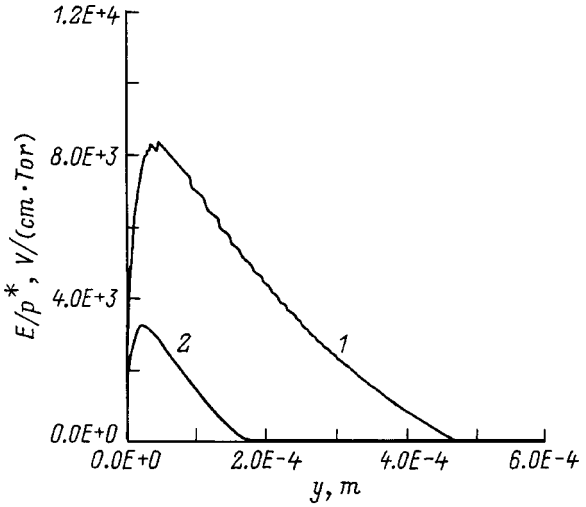

FIG. 2. The distribution of $(E/p^*)_K$ in the cathode fall region for $p^*=80$ Torr and $T_0=150$ (*1*), 350 K (*2*).

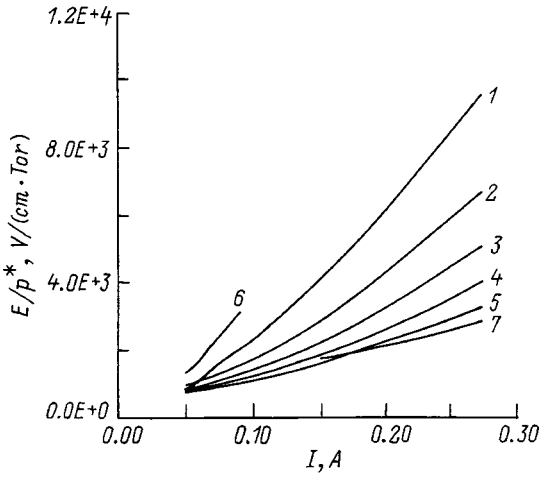

FIG. 3. The dependence of the maximum value of $(E/p^*)_K$ on the current at a cathode element for $p^*=80$ (*1–5*), 40 (*6*), and 160 Torr (*7*) and $T_0=150$ (*1*), 200 (*2*), 250 (*3*), 300 (*4*), 350 (*5,7*), and 120 K (*6*).

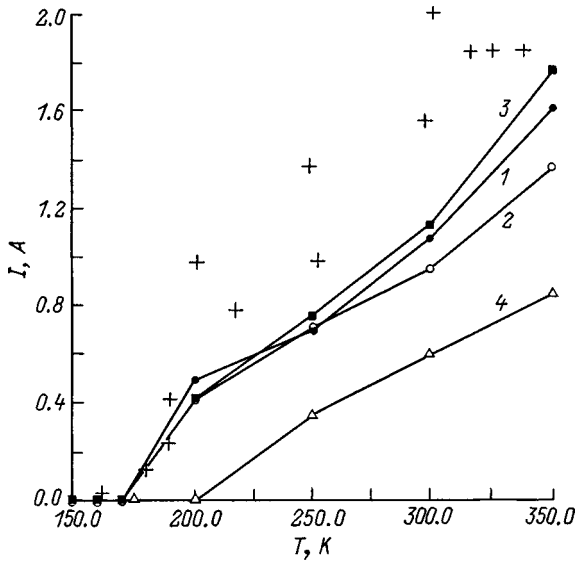

FIG. 4. The limiting discharge current as a function of gas temperature: + — experiment; *1,2,4* — laminar boundary layer; *3* — turbulent boundary layer; $H=0.03$ (*1–3*), 0.048 m (*4*); $L=0.1$ (*1,4*); 1 m (*2,3*).

Figure 3 shows plots of the maximum of $(E/p^*)_K$ as a function of the cathode current. The horizontal line denotes the critical level we have assumed. It was assumed that the inhomogeneity of the discharge is lost when the parameter $(E/p^*)_K$ at an individual cathode reaches a value corresponding to the maximum of $(E/p^*)_K$ for $T_0=350$ K and $I=242$ mA. It is evident from Fig. 3 that, as the initial temperature is lowered, the degree of anomaly of the cathode sheath increases and there is a significant drop in the critical current. At $T_0=160$ K the critical level of $(E/p^*)_K$ is reached by $I=122$ mA. This value is roughly equal to the current jump during breakdown of the discharge gap. In fact, given the difference between the breakdown and working electric field strengths, the current jump at the cathode during breakdown can be calculated using the formula

$$\Delta I_i \approx 0.32\left(\frac{E}{p^*}\right)\frac{Hp^*}{R_i}. \tag{18}$$

At a pressure of $p^*=80$ Torr it equals 123 mA. Thus, for temperatures below $T_0=160$ K, a uniform discharge was not ignited. The chosen critical level of $(E/p^*)_K$ is in agreement with other data from Ref. 4. For example, when p^* was increased from 40 to 160 Torr there,[4] the critical temperature rose from 120 to 350 K. The jumps in the breakdown current given by Eq. (18) for $p^*=40$ and 160 Torr are equal to $\Delta I=61$ and 246 mA, respectively. It is evident from Fig. 3 that for $p^*=40$ Torr and $T_0=120$ K and for $p^*=160$ Torr and $T_0=350$ K, the critical cathode currents equal 82 and 267 mA, respectively; that is, they are quite close to the current jumps during breakdown of the discharge gap.

Figure 4 shows the calculated and experimental[4] values of the maximum total discharge current for which a uniform discharge existed. It is obvious that, as in the experiment, increasing the initial gas temperature causes the limiting discharge current to rise. At the same time the current rose, the number of cathodes on which the gas broke down with a uniform discharge also increased. While the discharge did not ''ignite'' on any one of the cathodes at $T_0=150$ K, it burned from all ten at $T_0=350$ K. The calculated limiting currents are close to the level of the lower boundary of the experimental values.[4] The calculations also showed that the discharge stability depends on the flow parameters and the geometric dimensions of the channel. When the length of the segment of the channel attached to the discharge chamber was increased from $L=0.1$ to 1.0 m, the limiting discharge current was lowered. The transition from a laminar flow in the boundary layer to a turbulent flow increased the limiting discharge current slightly. Increasing the channel height from 0.03 to 0.048 m while maintaining the initial density and velocity of the gas reduces the limiting current substantially, by almost a factor of two, and increased the critical temperature. This is consistent with the results of Ref. 26.

If we had assumed that the mechanism for the discharge instability is purely based on field emission, then the match between the calculated and experimental data would have been considerably worse, since under these conditions the field strength at the cathode surface depended weakly on the current and was practically independent of the initial gas temperature. We have related the transition from a uniform to a nonuniform discharge to a level of $(E/p^*)_K=2734$ V/(cm·Torr) in the cathode fall region. If the assumption that a critical level of $(E/p^*)_K$ exists is fundamentally true, then its absolute magnitude may depend on the specific conditions.

CONCLUSIONS

A modification of the main formulas from the Engel–Steenbeck theory has made it possible to study the effect of a nonuniform distribution of the gas density on the parameters of the cathode dark space. The effect of a density nonuniformity increases as the pressure rises and is substantial for pressures exceeding 5 Torr. Calculations with a change in the initial gas temperature have shown that, for a given cathode current, the overheating of the gas in the cathode dark space and its degree of anomaly increase as the initial gas

temperature is lowered. In a transverse flow, as the gas temperature at the channel inlet is reduced, there is a increase in the nonuniformity in the distribution of the discharge current over the cathodes. These processes control the development of a critical gas temperature below which a uniform discharge is not ignited. The distribution of the current over the cathodes and the current density in the region of the cathode fall can determine the way the limiting discharge current depends on the gas dynamic structure of the flow and on the channel parameters. These characteristics can be related to the effect on the discharge stability of the gas flow regime in the flow boundary regions (laminar, turbulent), the channel height, and the length of the channel lying in front of the discharge chamber.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 95-01-00619).

[1] Yu. D. Korolev and G. A. Mesyats, *Physics of Pulsed Breakdown in Gases* [in Russian], Nauka, Moscow (1991), 224 pp.
[2] Yu. S. Akishev, A. P. Napartovich, S. V. Pashkin *et al.*, Teplofiz. Vys. Temp. **22**, 201 (1984).
[3] G. A. Baranov and S. A. Smirnov, Zh. Tekh. Fiz. **69**(11), 42 (1999) [Tech. Phys. **44**, 1298 (1999)].
[4] S. V. Pashkin and P. I. Peretyat'ko, Kvantovaya Élektron. (Moscow) **5**, 1159 (1978) [Sov. J. Quantum Electron. **8**, 665 (1978)].
[5] W. S. Boyle and F. E. Howorth, Can. J. Phys. **24**, 1466 (1956).
[6] V. A. Shveĭgert and I. V. Shveĭgert, Teplofiz. Vys. Temp. **27**, 23 (1989).
[7] Yu. P. Raizer and M. N. Shneĭder, Fiz. Plazmy **15**, 319 (1989) [Sov. J. Plasma Phys. **15**, 184 (1989)].
[8] Yu. P. Raizer and M. N. Shneĭder, Teplofiz. Vys. Temp. **29**, 1041 (1991).
[9] S. Ya. Bronin and V. M. Kolobov, *Abstracts of the VI All-Union Conferemce on the Physics of Low-Temperature Plasmas* [in Russian] (1989), Vol. II, pp. 149–151.
[10] B. Ya. Moĭzhes and V. A. Nemchinskiĭ, Zh. Tekh. Fiz. **59**(4), 22 (1989) [Sov. Phys. Tech. Phys. **44**, 395 (1989)].
[11] G. V. Gadiyak, K. A. Nasyrov, V. A. Shveĭgert, and O. U. Uuémaa, ITPM Preprint No. 30-85, Institute of Theoretical and Applied Mechanics, Academy of Sciences of the USSR, Novosibirsk (1985).
[12] G. V. Gadiyak, V. A. Shveĭgert, and O. U. Uuémaa, Izv. Sib. Otdel. Akad. Nauk SSSR **21**(6), 41 (1988).
[13] A. von Engel and M. Steenbeck, *Elektrische Gasentladungen, Ihre Physik und Technik*, Vols. 1 and 2 [Springer, Berlin (1932, 1934); ONTI, Moscow–Leningrad (1936)].
[14] Yu. P. Raizer, *Physics of Gas Discharges* [in Russian], Nauka, Moscow (1987), 592 pp.
[15] S. Ya. Bronin, V. M. Kolobov, V. I. Sushkov *et al.*, Teplofiz. Vys. Temp. **18**, 46 (1980).
[16] A. N. Lobanov, Ya. I. Londer, L. P. Menakhin *et al.*, Zh. Tekh. Fiz. **52**, 1959 (1982) [Sov. Phys. Tech. Phys. **27**, 1204 (1982)].
[17] V. M. Fomin, I. V. Shaimova, and V. A. Shveĭgert, *Abstracts of the II All-Union Conference on the Physics of Electrical Breakdown in Gases* [in Russian], Tartu (1984), pp. 178–180.
[18] K. N. Ul'yanov and V. V. Chulkov, *ibid.*, pp. 172–173.
[19] S. A. Smirnov, *Materials from the Conference on Physics and Technology of Plasmas* [in Russian], Minsk (1994), Part 1, pp. 168–171.
[20] S. A. Smirnov, TsNIIATOMINFORM NIIÉFA Preprint No. P-0918 [in Russian], D. V. Efremov Scientific-Research Institute of Electrophysical Apparatus, Moscow (1993), 64 pp.
[21] E. V. Gracheva, K. I. Kolchin, E. F. Prozorov *et al.*, Teplofiz. Vys. Temp. **26**, 644 (1988).
[22] J. Meek and J. Craggs, *Electrical Breakdown of Gases* [Clarendon Press, Oxford (1953); IL, Moscow (1960)].
[23] Yu. M. Belyakov, F. M. Gaĭsin, G. Yu. Dautov *et al.*, Teplofiz. Vys. Temp. **24**, 802 (1986).
[24] Yu. S. Akishev, K. V. Baiadze, V. M. Vetsko *et al.*, Fiz. Plazmy **11**, 999 (1985) [Sov. J. Plasma Phys. **11**, 582 (1985)].
[25] E. P. Velikhov, V. S. Golubev, and S. V. Pashkin, Usp. Fiz. Nauk **137**, 117 (1982) [Sov. Phys. Usp. **25**, 340 (1982)].
[26] A. V. Artamonov, V. I. Blokhin, A. A. Vedenov *et al.*, Kvantovaya Élektron. (Moscow) **4**, 581 (1977) [Sov. J. Quantum Electron. **7**, 323 (1977)].

Translated by D. H. McNeill

Development of ionization in nonequilibrium inert-gas plasmas in magnetogasdynamic channels

R. V. Vasil'eva, E. D. D'yakonova, A. V. Erofeev, and T. A. Lapushkina

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted June 5, 1998)
Zh. Tekh. Fiz. **69**, 56–61 (November 1999)

Effects accompanying the interaction of a flow of preionized inert gas with a magnetic field are studied: selective electron heating, the development of nonequilibrium ionization, and the onset of the ionization instability. Local and average densities and temperatures of the electrons are measured and the average ionization rate is determined. It is found that the average electron density increases as the magnetic induction is raised, in both stable and ionization unstable plasmas. The difference in the rates at which ionization develops in these two states is revealed. The mechanism for the coupling between the average ionization rate in an ionization unstable plasma and the spatial–temporal characteristics of the plasma inhomogeneities is established. © *1999 American Institute of Physics.* [S1063-7842(99)01011-9]

This paper is a continuation of a series of papers on the possibility of using pure inert gases as working materials in nonequilibrium magnetogasdynamic channels.[1–5] These studies have been made at magnetic fields below the critical values for development of the ionization instability, as well as at fields above critical. Special attention has been devoted to studying the development of ionization and the ionization instability in the plasmas. It should be noted that most of the studies of the ionization instability have been done in inert gases with an alkali metal additive.[6] In this case, the ionization of the gas is determined by the ionization of the additive, the degree of ionization in the initial state and in the inhomogeneities is close to the equilibrium values determined by the electron temperature, and the maximum electron density in the inhomogeneities is limited by complete ionization of the additive. The ionization processes in the pure inert gases differ in that, for the plasma parameters typical of magnetogasdynamic channels, the characteristic ionization and recombination times in them are orders of magnitude higher than in the alkali metals, so that the plasma state is far from equilibrium. In addition, there are no limits on the maximum electron density. Therein lies the novelty of these plasmas compared to those employed previously. It turns out that the ionization instability in pure inert-gas plasmas has an unusual property. It has been observed previously that their effective conductivity increases when the magnetic field is raised.[1–3] In the present study, primary attention is devoted to studying the structure of the inhomogeneities and their lifetime, to finding a correlation between the local temperature and density of the electrons in the inhomogeneities and their average values and to determining their average ionization rate.

The experiment was done in a disk magnetogasdynamic channel attached to a shock tube in which a flow of thermally equilibrium gas was created. An annular Faraday current was induced in the disk magnetogasdynamic channel, while the Hall current was not closed. The length of the interaction zone in the magnetogasdynamic channel was 0.1 m and the channel height was 0.01 m. The Mach number of the flow in the magnetogasdynamic channel lay in the interval 4–2, the magnetogasdynamic interaction parameter, 0–0.05, the Hall parameter, 0–4, and the conductivity, 100–600 S/m. The apparatus and measurement techniques have been described elsewhere.[1–4] The electron temperature was determined from the decay of the continuum in the near ultraviolet and the electron density, from the continuum intensity.[5] The parameters of the flow behind the incident shock wave in the shock tube, which are close to the parameters in the magnetogasdynamic channel, were used as a standard. The experiments were done in xenon in three gasdynamic regimes I, II, and III, specified by Mach numbers 6.4, 6.9, and 8.4, respectively, in the shock tube. The initial pressure in the low pressure chamber was 26 Torr. Most of the experiments were done in regime II. Figure 1 is a sketch of the apparatus.

The dynamics of the evolution of the luminous inhomogeneities (magnetic striations) was studied using a high speed moving picture camera which recorded fragments of the disk channel with the aid of rotating mirrors. An analysis of the moving pictures showed that the striations have the form of spokes inclined at an angle of roughly 20°–30° to the azimuth, the azimuthal separation between them is about 90°, their radial length is about 0.03 m, and 2–3 of them exist simultaneously in the channel. The striations develop with a certain spacing and their propagation velocity is close to the flow velocity. As the striations propagate along the channel, their luminosity increases, they expand somewhat, they orient themselves more closely to the direction of the initial current, and their structure becomes more complicated. Figure 2 shows photometric traces of a film along the channel radius. It should be noted that the ordinate is the relative photometer reading, for which the maximum darkening exceeds the linear darkening limits for the film. Here it is possible to follow the appearance of individual striations and their increasing luminosity as they propagate along the

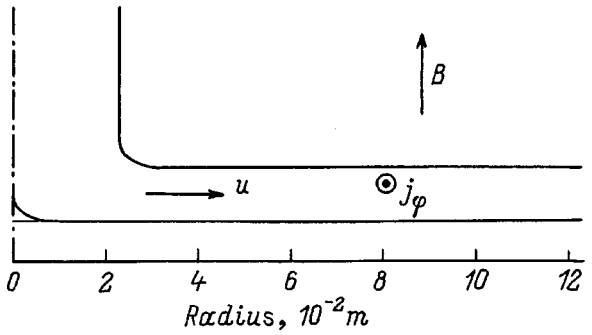

FIG. 1. Sketch of the magnetogasdynamic channel.

channel. Once they appear, the striations do not decay. Their lifetime exceeds the drift time. The evolution of the striations can be followed in more detail by analyzing oscilloscope traces of the light detected by photomultipliers located at two radii r_1 and r_2 at a single azimuth. Figure 3 shows the time variation in the electron densities at two radii in regime II as the different plasma volumes move past the observation window. The rise in the average electron density with time is caused by the finite ionization relaxation time in the plug of shock compressed gas. In the density fluctuations, large scale inhomogeneities can be identified and, against the background of these, smaller scale inhomogeneities that show up more clearly at the larger radius. Over a drift time of 4.5 $\times 10^{-4}$ s, about 8 large scale striations pass by. Given that $\Delta r = u\Delta t$ and the flow velocity is about 10^3 m/s, we estimate the transverse size of the large scale striations to be 0.02–0.3 m. The large scale striations follow one another by $3 \times 10^{-5} - 5 \times 10^{-5}$ s and the time between the appearance of the striations at the first and second cross sections is close to the time for them to move a distance $(r_2 - r_1)$. Small scale inhomogeneities with sizes of about 0.01 m or less appear toward the end of the channel and develop in the large scale inhomogeneities with a higher electron density. Here there is some analogy with the development of classical turbulence,[7] where small scale fluctuations are regarded as a fine detail structure imposed on a large scale inhomogeneity. As they move, the electron density in the magnetic striations increases.

Figure 4 shows the electron density and temperature in the fluctuations passing the observation port at $r = 0.096$ m for $B = 1$ T. The middle of the slug of hot gas has been isolated here. Evidently, an elevated electron temperature is observed mainly in regions with a higher temperature, and vice versa: in regions with a reduced electron density the temperature is also lower. In the large scale inhomogeneities the maximum an minimum electron densities differ by roughly a factor of 3–3.5, while the electron temperature varies from 8500 to 7000 K. In the small scale inhomogeneities, the electron density varies less, by roughly 50%, while

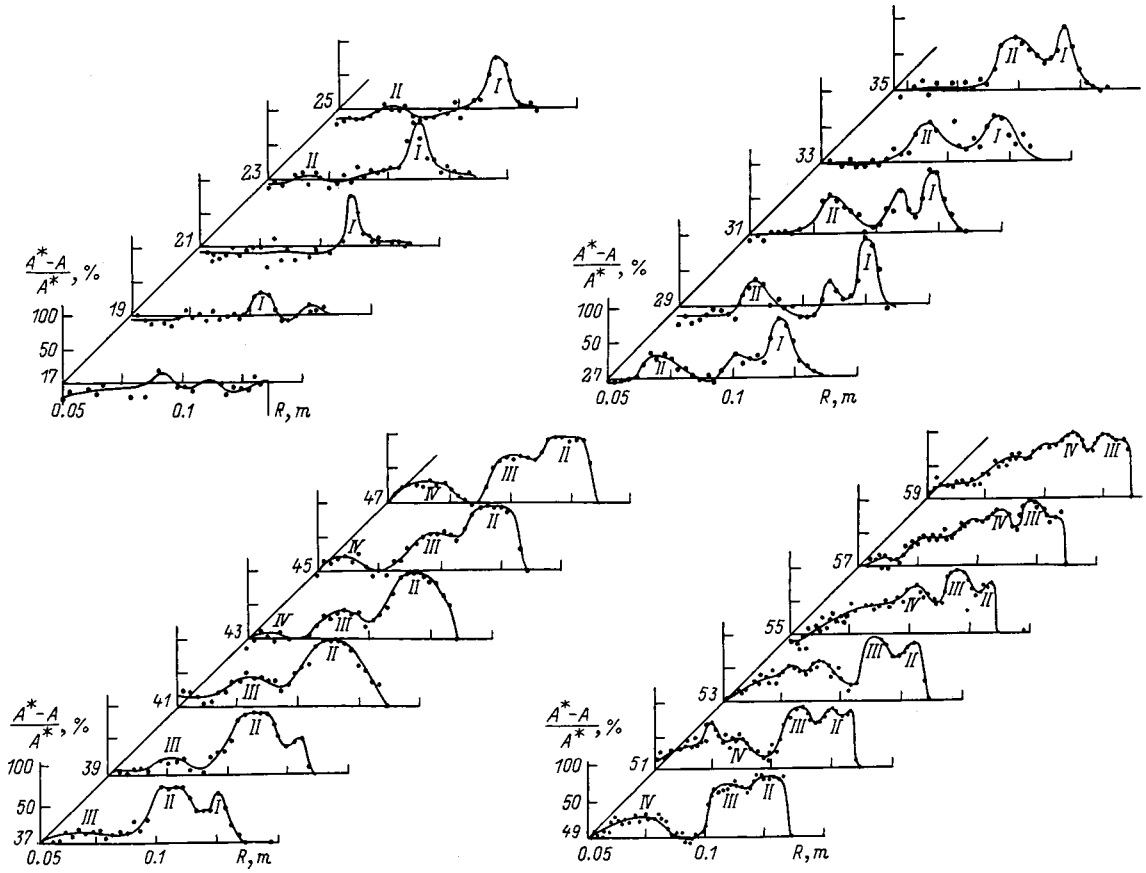

FIG. 2. Radial photometric scans of the moving picture frames. A is the instantaneous readout of the microphotometer and A^* is the background readout. The numbers on the scans denote the number of the picture frame. The interval is 3.2 μs. The Roman numerals represent the numbers of the striations which appear. Regime II. $B = 1T$.

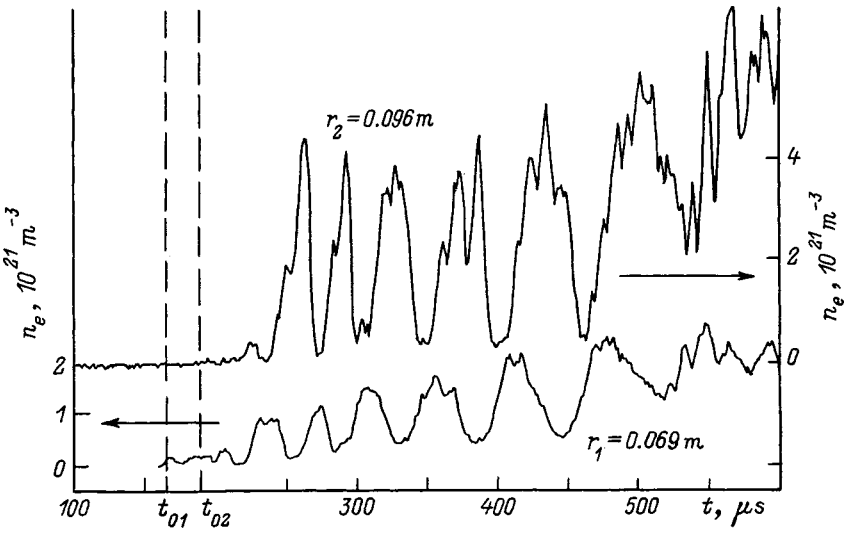

FIG. 3. The time variation in the electron density at two radii. Regime II. t_{01} and t_{02} are the arrival times of the shock front at the first and second measurement cross sections.

the electron temperature goes higher, to 10 000–11 000 K.

The electron density rose little, probably because the small scale inhomogeneities developed later and their lifetime is insufficient for ionization to develop strongly. It should be noted that the lowest values of T_e exceed its values at the channel inlet.

Figure 5 shows the electron temperature, averaged over the channel, in regime II. The values of $\langle T_e \rangle$ at the two radii r_1 and r_2 differ little from one another. Also shown here is $T_e = f(B)$ calculated assuming a stable plasma. This curve illustrates the features of the process at a collisional level. $T_e = f(B)$ is nonmonotonic and as B is increased, T_e decreases, since Coulomb collisions play an increasing role as ionization develops and this leads to a higher rate of energy transfer, and, therefore, to a reduction in the selective heating of the electrons. In addition, selective electron heating is affected by a slight reduction in the flow velocity owing to the ponderomotive force. At the average of the radii examined here ($r = 0.085$ m), when B increases from 0 to 1 T, the flow velocity decreases from 1.2×10^3 to 1.0×10^3 m/s and the calculated densities of the atoms increase from 1.0×10^{24} to 1.2×10^{24} m^{-3}. The calculated temperature of the heavy component increases more significantly, from 1800 to 5500 K. Measurements of T_e in the stability region indicate good agreement with the calculations. The average values of T_e in the ionization unstable plasma are lower than those calculated assuming stability. It should be noted that previous experiments showed that the effective conductivity of the ionization unstable plasma increases with the magnetic field.[1,3] This ought to reduce to more intense Joule heating of the electrons, but this probably does not happen because part of the Joule energy is lost in fluctuations.[8]

The degree to which the instability develops, which is expressed in the relative fluctuations in n_e, depends on the degree of supercriticality, B/B_{cr}, of the magnetic field. Figure 6 shows the root mean square relative fluctuations in n_e as a function of B in the three regimes. B_{cr} is determined from the sharp increase in fluctuations; the larger B is compared to B_{cr}, the higher the level of fluctuations is. For a given radius, B_{cr} increases from regime to regime. This happens because of the difference in the degree of ionization of the gas at the inlet to the magnetogasdynamic channel: $\alpha_{0I} = 10^{-4}$, $\alpha_{0II} = 3\times10^{-4}$, $\alpha_{0III} = 1.1\times10^{-3}$. For high degrees of ionization, because of the increased role of Coulomb collisions there is an increased rate of momentum transfer, which leads to less selective heating of the electrons, increases the characteristic ionization time and, therefore, in-

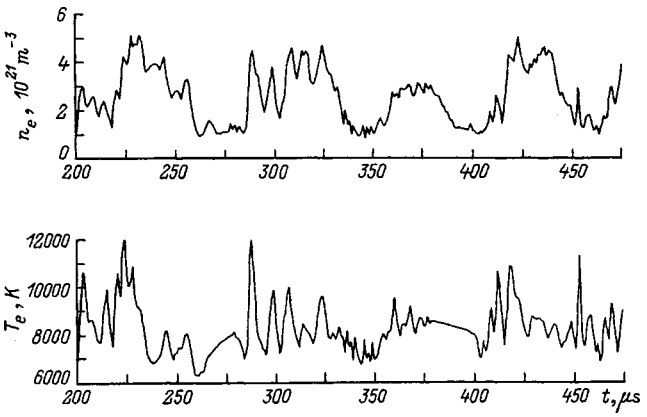

FIG. 4. Variations of the electron density and temperature in inhomogeneities at $r_2 = 0.096$ m. Regime II.

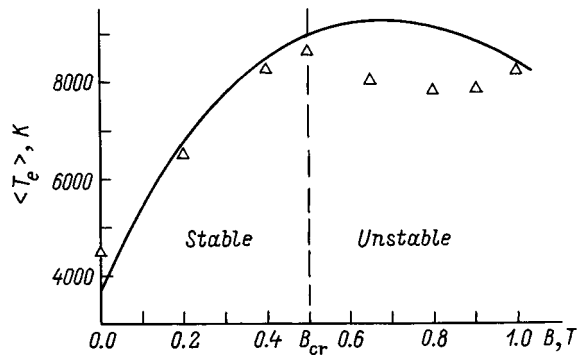

FIG. 5. Average electron temperature as a function of magnetic induction. The points are experimental data and the curve was calculated assuming a stable plasma.

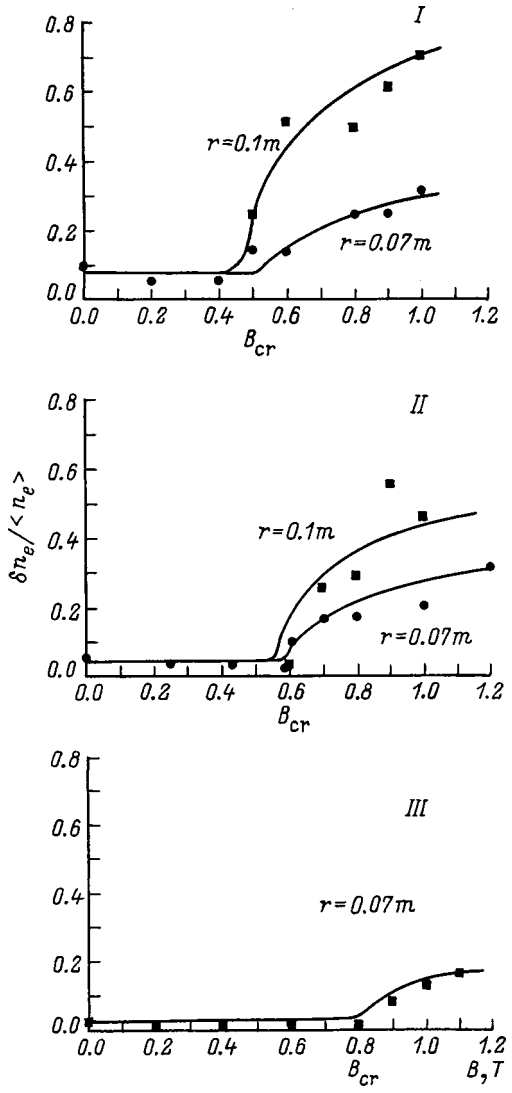

FIG. 6. Relative fluctuations in the electron density as a function the magnetic field in regimes I–III.

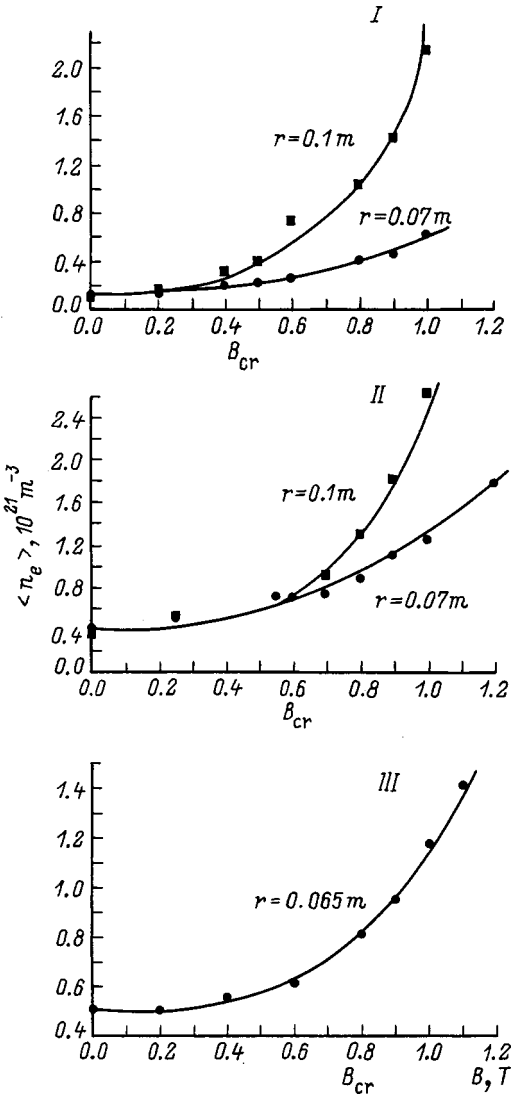

FIG. 8. Average electron density as a function of magnetic induction in regimes I–III.

creases the time for the instability to develop. Thus, in order for the instability to develop at a given radius, the magnetic field must increase. B_{cr} at different radii was determined in more detail in regime II. The results are plotted in Fig. 7. When the magnetic field is higher, the instabilities develop closer to the channel inlet, their lifetime in the magnetogasdynamic channel is longer, and the fluctuations observed at the end of the channel are stronger.

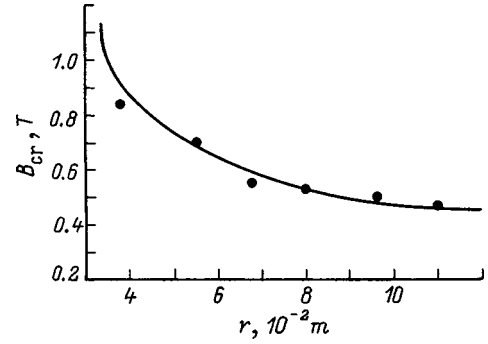

FIG. 7. Critical magnetic field as a function of channel radius. Regime II.

Figure 8 shows the electron density, averaged over the channel, as a function of the magnetic field at different radii for the three regimes. Each point was obtained by analyzing several oscilloscope traces. Evidently, $\langle n_e \rangle$ is higher for higher B and increases along the radius. $\langle n_e \rangle$ increases the most in regime I, where the degree of supercriticality of the magnetic field is higher and the least in regime III, where the degree of supercriticality is lower. In all cases, the electron densities are substantially lower than the equilibrium values corresponding to the average electron temperatures.

The average ionization rates were estimated by analyzing several measurements of $\langle T_e \rangle$ and $\langle n_e \rangle$ at the two radii, $r_1 = 0.069$ m and, $r_2 = 0.096$ m,

$$\frac{\Delta\langle n_e\rangle}{\Delta t} = \frac{\langle n_{e2}\rangle - \langle n_{e1}\rangle}{\Delta t} - \frac{\langle n_{e1}\rangle}{n_{a1}} \cdot \frac{n_{a2} - n_{a1}}{\Delta t}.$$

The second term accounts for small corrections associated with the change in the density of atoms and is taken

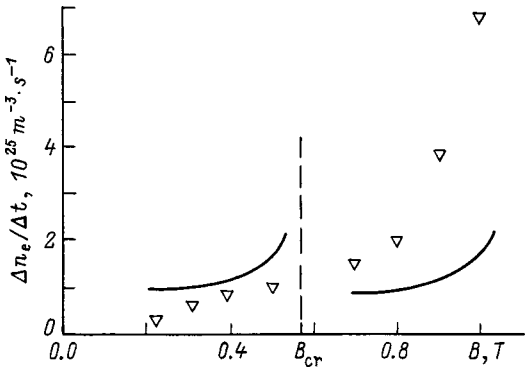

FIG. 9. Average ionization rate as a function of magnetic induction. Regime II. The points are experimental data and the curves were calculated.

from calculated data. Since the flow velocity (u) varies little in this segment of the channel, $\Delta t=\langle r_2-r_1\rangle/u$. The ionization rates measured in this way are comparable to the calculated values for the average temperatures and densities of the electrons,

$$\frac{\Delta\langle n_e\rangle}{\Delta t}=\frac{\langle n_{e1}\rangle}{\Delta t}[\exp(t_{\rm fl}/\langle\tau_i\rangle)-1],$$

where

$$t_{\rm fl}=(r_2-r_1)/u, \quad \langle\tau_i\rangle=\frac{1}{K_i\cdot\langle T_e\rangle\cdot\langle n_a\rangle},$$

$t_{\rm fl}$ is the drift time (time of flight) and τ_i is the ionization time.

The values of the ionization coefficient, K_i, were taken from Ref. 9. Recombination plays no significant role, so it was neglected. The results of this analysis are shown in Fig. 9. In the stable plasma region with $B<B_{\rm cr}$, the measured ionization rates were lower than the calculations. However, the deviation lies within the experimental error, since, in this region, Δn_e is given by a small difference of large quantities and the accuracy is no better than 50%; an error of 3% in T_e gives an error of about 50% in the ionization coefficient. In the ionization unstable plasma region, $B>B_{\rm cr}$, the average ionization rate is, on the other hand, considerably higher than the calculations for the average temperatures. This happens primarily because regions with elevated temperatures play a dominant role in the development of ionization through raising the average electron density.

The overall picture of the instability development looks like the following. Fluctuations in the plasma electron temperature develop and the degree of ionization changes in accordance with the kinetics. In layers with an elevated temperature, which are inclined at a certain angle to the initial current because of the additional currents created by the electron density gradient in a transverse magnetic field, more power is released and this may exceed the increased electron energy loss in collisions with heavy plasma species. This, in turn, increases the electron temperature and the degree of ionization. Thus, oscillations are driven. A linear analysis[1] shows that in an ionized plasma where the degree of ionization of the gas is much lower than the equilibrium level, the ionization instability can develop for arbitrarily low values of the Hall parameter, but the lower the Hall parameter is, the longer the time for the ionization instability to develop will be. The time for the instability to develop is comparable in order of magnitude to the characteristic ionization time. Thus, the concept of a critical magnetic field, as this experiment has shown, is meaningful only for a certain radius of the magnetogasdynamic channel and means that, in this channel, the selective heating of the electrons is sufficient to make the characteristic ionization time shorter than the drift time. An analysis of the experimental data shows that the nonlinear stage of the instability development proceeds in this fashion. Initially, as the plasma loses stability with respect to small oscillations in the electron temperature and, therefore, the density, large scale regions with elevated and reduced electron temperatures develop in the plasma. The average electron temperatures are determined by the electron energy balance and greatly exceed the gas temperature. It turns out that these inhomogeneities are long lived objects, at least their lifetime exceeds the drift time. In a first approximation we can assume that volumes of plasma are somehow ''frozen'' into these regions. With the passage of time, in the hotter regions (striations) as the electron temperature oscillations are pumped more strongly, the electron density increases. In the regions with a reduced temperature, the electron density does not change significantly because of the long characteristic recombination time. Thus, the average conductivities in the plasma increase. As large scale inhomogeneities develop at the end of the channel, bifurcation of the plasma state takes place and small scale inhomogeneities with higher electron temperatures show up against the background of the large scale striations. Probably, if the magnetic field is increased further or the channel length is increased, the striations will be destroyed and an irregular, turbulent distribution of the electron density and currents will develop. The results in this paper apply to this phase of the development of the ionization instability, when regular structures exist within it. Later, when the striations have an elongated shape in the form of spokes and approach the direction of the Faraday current, an increase in the average conductivity, despite the stratification of the flow, will produce an increase in the effective conductivity and, accordingly, in the Joule heating. This explains the experimentally observed rise in the electron density on moving along the radius and in the average electron density as the magnetic induction is increased.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 98-01-01121).

[1] R. V. Vasil'eva, A. L. Genkin, V. L. Goryachev *et al.*, *Nonequilibrium Inert-Gas Plasmas with Nonequilibrium Ionization in Magnetogasdynamic Generators* [in Russian], A. F. Ioffe Physicotechnical Institute, St. Petersburg (1991), p. 206.
[2] R. V. Vasil'eva, E. A. D'yakonova, A. V. Erofeev *et al.*, Zh. Tekh. Fiz. **67**(12), 6 (1997) [Tech. Phys. **42**, 1376 (1997)].
[3] T. A. Lapushkina, R. V. Vasil'eva, A. V. Erofeev, and A. D. Zuev, Zh. Tekh. Fiz. **67**(12), 12 (1997) [Tech. Phys. **42**, 1382 (1997)].
[4] A. V. Erofeev, R. V. Vasil'eva, A. D. Zuev *et al.*, in *Proceedings of the 12th International Conference on MHD Electrical Power Generation*, Yokohama, Japan (1996), Vol. 1, pp. 74–82.

[5] T. A. Lapushkina, E. A. D'yakonova, and R. V. Vasil'eva, Pis'ma Zh. Tekh. Fiz. **24**(2), 58 (1998) [Tech. Phys. Lett. **24**(1), 66 (1998)].
[6] A. V. Nedospasov and V. D. Khait, *Elements of the Physics of Processes in Low-Temperature Plasma Devices* [in Russian] (1991), 224 pp.
[7] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* [Pergamon Press, Oxford (1959); GITTL, Moscow (1954), 796 pp.].
[8] R. V. Vasil'eva, E. A. D'yakonova, A. V. Erofeev, and T. A. Lapushkina, Pis'ma Zh. Tekh. Fiz. **24**(14), 48 (1998) [Tech. Phys. Lett. **24**(7), 562 (1998)].
[9] A. M. Biberman, V. S. Vorob'ev, and I. T. Yakubov, Teplofiz. Vys. Temp. **7**, 593 (1969).

Translated by D. H. McNeill

Emission properties of a plasma cathode based on a glow discharge for generating nanosecond electron beams

V. I. Gushenets, N. N. Koval', V. S. Tolkachev, and P. M. Shchanin

Institute of High Current Electronics, Siberian Branch of the Russian Academy of Sciences, 634055 Tomsk, Russia

(Submitted July 27, 1998)
Zh. Tekh. Fiz. **69**, 62–65 (November 1999)

The emission properties of a plasma cathode based on a nanosecond pulsed glow discharge with currents of up to 200 A at a pressure of 5×10^{-2} Pa are studied experimentally. Stable ignition and burning of the discharge are ensured if the current in the auxiliary pulsed discharge is 25–30% of that in the main discharge and its pulse duration exceeds that of the main discharge by more than an order of magnitude. Emission current pulses from the cathode with amplitudes of up to 140 A fully reproduce the discharge current and are determined by the transparency of the grid anode. © *1999 American Institute of Physics.* [S1063-7842(99)01111-3]

1. INTRODUCTION

High current, nanosecond electron beams with short current pulse rise and fall times are required for electron and x-ray pumping of lasers. Usually accelerators with explosive emission cathodes, which provide currents of up to a few kiloamperes,[1,2] or accelerators with plasma cathodes based on vacuum arcs[3–5] are used for these purposes In both cases, the continuous operating period is limited by the lifetime of the cold cathode to on the order of 10^7 pulses, at best. Furthermore, when the beam area is large, it is difficult to ensure a uniform distribution of the current density over the beam cross section. A long continuous operating time can be obtained by using a hollow cathode glow discharge with pulsed current densities of up to 1 A/cm^2 as an emitter.[6,7] The high gas pressure (1–10 Pa) required for stable ignition and burning of a discharge at high currents and the use of He as a working gas at high accelerator voltages on the order of 100 kV limits the use of these discharges only in dc accelerators with current densities of several tens of milliamperes.[8,9] The operating pressure in a hollow cathode glow discharge has been lowered to 10^{-2} Pa by operating under conditions such that the cathode area (S_k) exceeded the anode area (S_a) by $\sqrt{M/m}$ times.[10] Thus, when nitrogen is used as a working gas, the cathode area should be at least 200 times the anode area. With a large beam cross sectional area, this kind of discharge system creates design problems associated with the large cathode size. Additional shortcomings of this system, which show up during the shaping of nanosecond pulses, will be discussed later in this paper.

2. EXPERIMENTAL SETUP AND RESULTS

The studies discussed below were aimed at clarifying the possibility of creating an accelerator that is ultimately capable of generating a wide aperture electron beam with an energy of hundreds of keV, a current of several hundred amperes with nanosecond duration, a high pulse repetition rate, and a lifetime on the order of 10^8-10^9 pulses. Because of the large size of the cathode, a discharge system that satisfies the condition $S_k/S_a=(M/m)^{1/2}$ was not studied. Initially, the emission properties of a standard discharge system, similar to that of Ref. 9, which consisted of a hollow cathode with a diameter of 20 cm and a length of 75 cm, a mesh anode of length 70 cm and width 3 cm located a distance of 5 mm from a rectangular window of area 2 $\times$70 cm in the side surface of the cathode, an auxiliary anode in the form of a rod with a diameter of 2 mm, and an extraction electrode mounted a distance of 1 cm from the mesh anode.

In this system, a low current initiating discharge is ignited between the auxiliary anode and the hollow cathode with a current of up to 5 mA in an He pressure above 1 Pa, while the main pulsed discharge is ignited between the same hollow cathode and the mesh anode. Two burning regimes for the pulsed discharge were observed: normal and anomalous with strong oscillations in the discharge current. Similar burning regimes have been observed in hollow cathode discharges before.[10] Figures 1 and 2 show the current–voltage characteristic and the pressure dependence of the discharge current with a constant voltage on the discharge gap for the normal discharge. The transition from one burning regime to the other depends on the pressure of the working gas and on the rate of rise and amplitude of the discharge current.

It was also found that when the pressure is reduced, the time to form the discharge increases. The discharge current rises to its maximum over a time of 20 μs at a pressure of $P=3.9$ Pa and over 45 μs at $P=1.4$ Pa. The discharge formation time and, therefore, the duration of the discharge current pulse, can be reduced by creating a substantial overvoltage across the discharge gap. However, this increases the rate of rise of the current, which causes the discharge to undergo a transition into the anomalous regime with a strong modulation in the current, as noted above, initially in the front. As the pulse duration is reduced or the discharge current increases further, the amplitude of the modulations increases and gradually the entire pulse is modulated by high

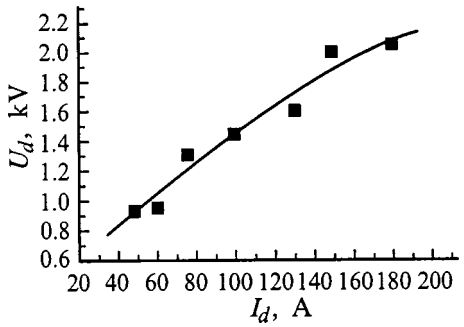

FIG. 1. Current–voltage characteristic of the hollow cathode discharge.

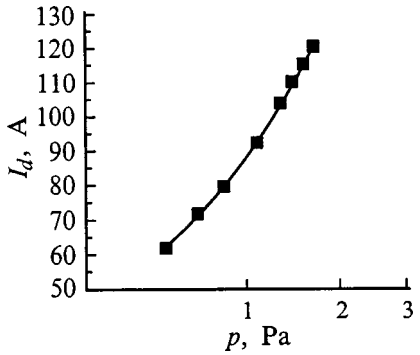

FIG. 2. The hollow cathode discharge current as a function gas pressure (helium).

frequency oscillations. In addition, in the experiments we observed contraction of the main discharge, which could be avoided by turning to the use of a low current dc discharge for initiating the high current pulsed discharge. Electron extraction experiments showed that for an accelerating voltage of 17 kV applied between the mesh anode and extraction electrode, the plasma cathode yielded an emission current of up to 100 A with a discharge current of 120 A and a pressure of 2.6 Pa. The current pulse shape was sinusoidal with a duration of 5 μs at its base.

The above shortcomings, as well as the high gas pressure, narrow range of working pressures for which the discharge operates stably, and low electrical breakdown strength of the accelerating gap at high pressures have stimulated a study of a discharge system in which a low gas pressure is attained in the hollow cathode and accelerating gap through preliminary filling of the cathode hollow with an auxiliary pulsed discharge plasma.

The experimental apparatus is shown schematically in Fig. 3 and consists of a hollow cathode with a diameter of 20 cm and length of 75 cm with two plasma generators mounted on its ends to create a preplasma in the hollow cathode. The plasma generators are attached to the hollow cathode through small (diameter 5 mm) apertures across which a pressure drop is created such that the pressure in the plasma generators is an order of magnitude higher than in the hollow cathode. The working pressure in the cathode hollow usually did not exceed 5×10^{-2} Pa and nitrogen or air were used as the plasma forming gas. The plasma generator, in turn, consists of a hollow cathode *1* and an intermediate electrode *2*. In order to increase the lifetime of the system a glow discharge is also used in the plasma generators. A 0.3 T magnetic field created by annular permanent magnets reduces the ignition voltage for the discharge, the working pressure in the plasma generators, and the jitter in the discharge ignition time. Oscilloscope traces of the initiating discharge between the cathode *1* and the electrode *2* and of the auxiliary discharge (used to fill the cathode hollow with plasma) between the intermediate electrode *2* and the hollow cathode *3* are shown in Figs. 4a and 4b, respectively.

The current in the initiating discharge and the delay in turning on the auxiliary discharge were chosen subject to the condition of filling the hollow anode uniformly with plasma from the auxiliary discharge. For shorter delays and lower currents in the initiating discharge, the plasma in the hollow cathode developed in the shape of a filament with temporal and spatial instability. The main nanosecond discharge voltage was applied between the hollow cathode and the mesh anode with a delay of 6–10 μs relative to the time the auxiliary discharge was ignited. The trigger and burning voltage for the main discharge lay within 4–6 kV.

The rate of rise of the current in the main discharge is determined initially by the shape of the applied voltage, but at the peak it depends on the delay in turning on the main

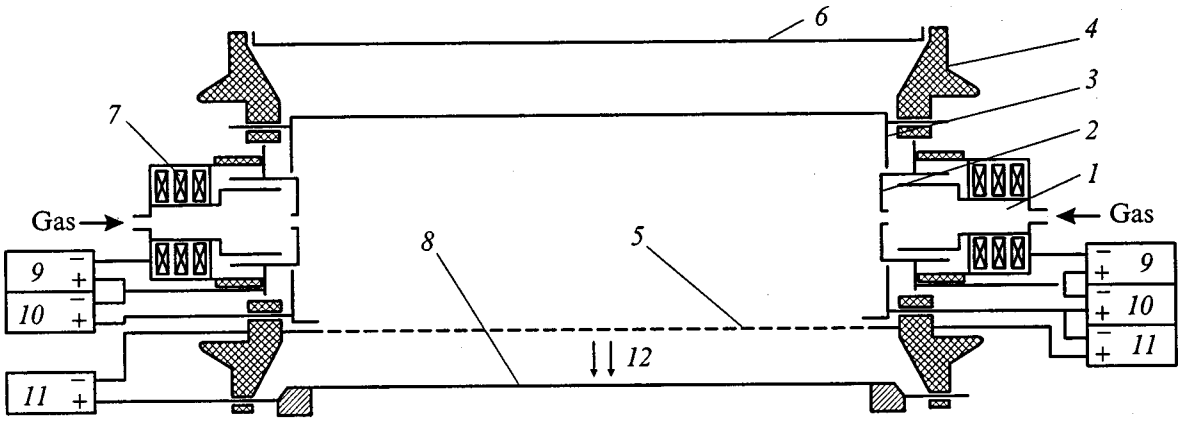

FIG. 3. Sketch of the experimental apparatus: *1* — plasma generator cathode, *2* — intermediate electrode, *3* — hollow cathode, *4* — insulator, *5* — mesh anode, *6* — vacuum vessel, *7* — magnets, *8* — extraction electrode, *9* — power supply PS1, *10* — PS2, *11* — PS3, *12* — electron beam.

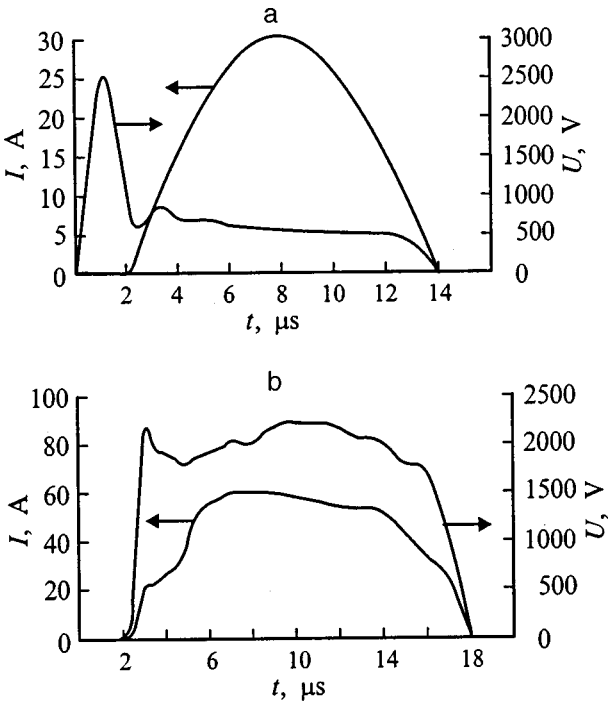

FIG. 4. Oscilloscope traces of the burning voltage and current of the initiating (a) and auxiliary discharges (b).

discharge relative to the time the auxiliary discharge is triggered, as can be seen in Fig. 5. Further reduction in the delay leads to a narrowing of the discharge current for fixed auxiliary discharge current and voltage between the hollow cathode and mesh anode. The shape of the emission current completely reproduces that of the discharge current. The maximum emission current obtained in these experiments was 140 A for an accelerating voltage of 15 kV between the mesh anode and extraction electrode. The electron extraction efficiency, which equals the ratio of the emission current to the discharge current, is proportional to the transparency of the mesh anode η, i.e., $I_e = \eta \times I_d$. The uniformity of the distribution of the current density along the major axis of the beam cross section was $\pm 15\%$.

Changing the polarity of the electrodes of the initiating and auxiliary discharges made it possible to reduce the burning voltage of the auxiliary discharge by a factor of 2, use

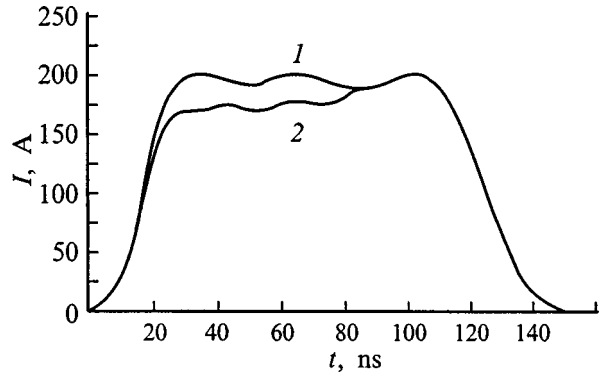

FIG. 5. Oscilloscope traces of the main discharge current (delay 10 (*1*), 6 μs (*2*)).

just one plasma generator for creating a plasma in the hollow cathode, substantially improve the emission current density distribution ($\pm 7\%$), and reduce the power demand by the plasma emitter by half.

In order to examine the continuous operating period of the plasma emitter we have tested the plasma generators, which determine the lifetime of the plasma emitter, at initiating and auxiliary discharge currents and voltages indicated above with a pulse repetition rate of 10^3 pulses per second for 4 h, which corresponded to 1.2×10^8 pulses. Visual examination revealed no significant changes in the configuration or sputtering of the copper electrode *1* or the stainless-steel electrode *2*.

3. DISCUSSION OF RESULTS

Even the use of a pulsed initiating discharge in a system of electrodes with a rod anode in a high current glow discharge does not solve all the problems of shaping nanosecond and microsecond pulses with short rise and fall times because of its low energy efficiency. At low pressures, the initiating discharge develops over times of a few tens of μs, and when the duration is reduced, strong oscillations of the type seen in Ref. 10 develop. We assume that the oscillations in the discharge current are caused by the time lag in the creation of charged particles and their low mobility in the plasma, rather than by explosive processes at the inner surface of the hollow cathode, as assumed in Ref. 10. Similar fluctuations occur in arc discharges with a hollow anode and a negative anode fall.[11]

Short-duration beams are formed in these systems either by introducing an additional control grid or by use of a pulsed accelerating voltage. For beam currents up to several hundred amperes and nanosecond current rise times, the capacitive currents exceed the beam currents, and this creates certain problems. In the system considered here, nanosecond beams are formed at low (4–5 kV) voltages and this greatly simplifies the accelerating system and substantially reduces the energy costs. Further acceleration of the electrons is done with a dc voltage applied to an accelerator gap.

The formation of filament discharges in a hollow anode at high discharge currents is probably related to the formation of a double layer in the aperture because of the large difference in the plasma pressures and densities. Electrons accelerated in the double layer make it easier for a high current to pass through the small aperture and ensure continuity of the current in the two plasmas with their different densities. The formation of a double layer can explain the high burning voltage of the auxiliary discharge. A higher burning voltage in a glow discharge has been observed[12] when a double layer develops.

As to the change in the pulse shape and the dependence of the peak current on the delay in switching on the main discharge, these can be explained as follows. For stable burning of a glow discharge with current I_d at low pressure, it is first necessary to create and maintain a certain plasma density in the hollow anode, in order to compensate the loss of

electrons from the hollow cathode into the anode region. When the density created by the auxiliary discharge is not high enough, the main discharge reacts by reducing the discharge current or by reducing the current at the beginning of the pulse. A comparison of the oscilloscope traces of Figs. 4b and 5 implies that the auxiliary discharge current should be 25–30% of the main discharge current. This value is consistent with the ratio of the currents in analogous dc glow discharge systems.[13]

4. CONCLUSION

1. Prefilling the hollow cathode of the glow discharge with an auxiliary discharge plasma makes it possible to reduce the gas pressure, avoid the use of He as a working gas, and reduce the probability of a discharge along long paths in the accelerating gap.

2. Replacing an arc discharge with a glow discharge makes it possible to extend substantially the continuous operation period for a plasma cathode at high currents and a high pulse repetition rate.

3. Generating nanosecond current pulses by forming an corresponding pulsed discharge at low voltages simplifies the accelerating voltage system.

[1] E. N. Abdullin, S. I. Gorbachev, A. M. Efremov *et al.*, Kvantovaya Élektron. (Moscow) **20**, 652 (1993) [Quantum Electron. **23**, 564 (1993)].
[2] E. N. Abdullin, S. P. Bugaev, S. I. Gorbachev *et al.*, in *Laser Optics 93. Proceedings of the SPIE Conference on Laser Physics*, St. Petersburg (1993), Vol. 2, pp. 38–45.
[3] V. I. Gushenets, N. N. Koval', D. L. Kuznetsov *et al.*, Pis'ma Zh. Tekh. Fiz. **17**(23), 26 (1991) [Sov. Tech. Phys. Lett. **17**(12), 834 (1991)].
[4] V. I. Gushenets and P. M. Shchanin, Zh. Tekh. Fiz. **66**(12), 25 (1993) [Sov. Phys. Tech. Phys. **41**, 1048 (1993)].
[5] S. W. A. Gielkens, P. J. M. Peters, W. J. Witteman *et al.*, Rev. Sci. Instrum. **67**, 2449 (1996).
[6] R. S. Knechtli and G. N. Mercer, IEEE J. Quantum Electron. **QE-9**, 684 (1973).
[7] D. M. Goebel, R. W. Schumacher, and R. M. Watkins, in *Proceedings of the IX International Conference on High-Power Particle Beams (Beams-92)*, Washington (1992), Vol. 2, pp. 1093–1098.
[8] Y. R. Bayless, R. S. Knechtli, and G. N. Mercer, IEEE J. Quantum Electron. **QE-10**, 213 (1974).
[9] Y. R. Bayless, Rev. Sci. Instrum. **46** 1158 (1975).
[10] A. S. Metel', Zh. Tekh. Fiz. **56**, 2329 (1986) [Sov. Phys. Tech. Phys. **31**, 1395 (1986)].
[11] R. Dollinger, D. R. Dettman, A. S. Gilmour *et al.*, IEEE Trans. Plasma Sci. **PS-8**, 302 (1980).
[12] A. S. Metel', Zh. Tekh. Fiz. **54**, 241 (1984) [Sov. Phys. Tech. Phys. **29**, 141 (1984)].
[13] A. V. Vizir', E. M. Oks, P. M. Shchanin, and G. Yu. Yushkov, Zh. Tekh. Fiz. **67**(6), 27 (1997) [Tech. Phys. **42**, 611 (1997)].

Translated by D. H. McNeill

Conductivity of nano-MIM diodes with a carbonaceous active medium in a model including percolation effects

V. M. Mordvintsev and V. L. Levin

Institute of Microelectronics, Russian Academy of Sciences, 150007 Yaroslavl, Russia
(Submitted June 20, 1996)
Zh. Tekh. Fiz. **69**, 66–73 (November 1999)

A model proposed previously for processes in nano-MIM (metal–insulator–metal) diodes with a carbonaceous active medium is developed and refined. The inclusion of percolation effects in the insulating gap yields qualitatively new results and provides better agreement between the calculations and experimental data for physically reasonable values of all the parameters. An analysis of the model has made it possible to distinguish two different elements in the mechanism upon which it is based, which are important for understanding the essence of the processes that take place in electroformed nano-MIM diodes with a carbonaceous active medium: an internal negative feedback in the structure and modulation of the parameters of the cathode potential barrier. These elements show up in different ways in the observed characteristics of MIM structures. © *1999 American Institute of Physics.* [S1063-7842(99)01211-8]

INTRODUCTION

The N-shaped current–voltage characteristics of diodes with a metal–insulator–metal structure (MIM-diodes) placed in a vacuum with the vapor of organic compounds and subjected to electroforming have been observed by many researchers over several decades.[1] Several attempts have been made at a not very deep theoretical description of these objects; of these, the most successful is evidently the model with multiple burned-out and newly regenerated conducting paths.[2] Based on experiments using a scanning tunneling microscope, a more detailed mechanism has been proposed[3] to explain the region of the current–voltage characteristic with a negative differential resistance, the important feature of which is a nanometer width of the insulating gap of the MIM structure that is open to the arrival of organic molecules. We refer to this kind of structure, as opposed to the traditional structures, by the term ''nano-MIM diode.''

This mechanism, referred to as barrier width modulation by stratification of the dielectric, includes the following processes. The current through the MIM structure is limited by tunneling of electrons through a barrier near the cathode; because of the nanometer size of the gap, the required electric fields are achieved even at low voltages. Passage of electrons injected from the cathode through a gap filled with organic molecules is accompanied by electron-impact dissociation of the latter, which leads to stratification of the organic dielectric owing to the release of particles in a conducting carbonaceous phase from the anode side. The corresponding region, which we refer to as a formed dielectric, is a composite material (''conducting particles in a dielectric matrix''). This changes the potential distribution in the gap, reducing the barrier width near the cathode, which leads to an increase in the current through the structure and to heating of the dielectric. Because of some thermally activated process, such as desorption, the concentration of particles in a conducting carbonaceous phase decreases with the temperature of the formed dielectric, which, on the other hand, should reduce the current. In is significant that this type of mechanism includes an internal negative feedback, with one of the links in the feedback loop being the structural characteristics of the system, specifically the concentration of the particles in the conducting carbonaceous phase and the barrier width that depends on it. Thus, we can speak of self-formation of a nanometer structure in the insulating gap. The occurrence of these nonlinear processes in the insulating gap of a MIM structure, which lead to an N-shaped current–voltage characteristic, is indicative of a carbonaceous active medium in the device.

There is convincing experimental evidence[4,5] that a nanometer segment of the insulating gap is actually operating in the case of traditional MIM structures, as well. Thus, the proposed mechanism may be universal. Furthermore, it has been shown[6] that electroforming essentially involves the self-organization of the nanometer insulating gap in the carbonaceous conducting medium formed in these devices. A model[7] based on this mechanism provides a qualitatively correct description of the current–voltage characteristics nano-MIM diodes for physically reasonable values of all the parameters but one. This parameter, the maximum concentration (volume fraction) of particles in the conducting carbonaceous phase in the formed dielectric, α, ends up too large. Here a percolation threshold must be reached,[8] i.e., a qualitatively new effect appears: part of the formed dielectric in the insulating gap becomes a conducting medium owing to percolation and this should change the situation significantly. In this paper we propose a simplified model for nano-MIM diodes that includes this effect.

MODEL

We shall follow Refs. 3 and 7 on the basic elements of the mechanism and the approximations to be used, except where specially noted. Here the goal is to describe the sta-

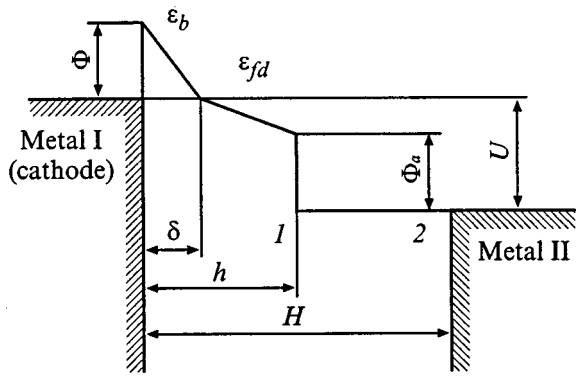

FIG. 1. Simplified potential diagram of an MIM structure with percolation in the formed dielectric: Φ is the barrier height for electrons at the cathode-insulating gap interface; Φ_a is the barrier height at the boundary of the formed dielectric (*1*) with the carbonaceous conducting medium (anode, *2*); U is the voltage across the electrodes; H is the width of the insulating gap of the MIM structure; and, h is the width of the insulating gap.

tionary situation in the gap of a flat MIM structure for different voltages U across its metallic electrodes, which lie a fixed distance H from one another in a single plane on the surface of an insulating substrate. Figure 1 shows a potential diagram for the structure.

As opposed to Ref. 7, we at once take the nonuniformity of the emission from the cathode surface into account (Fig. 2). Since emission is mainly from nanotips, the corresponding electron beams will diverge owing to defocusing in the local electric field. Scattering in the formed dielectric causes the electron flow to diverge further. It is rather difficult to take all these factors into account quantitatively, so we shall limit ourselves to introducing a constant divergence angle, 2ξ, for a beam of initial linear size r at the cathode and a uniform distribution of the current density in any transverse cross section. Given that the width of a beam in a plane perpendicular to the plane of Fig. 2 is constant (it is determined by the thickness of the film of formed dielectric), we obtain the following expression for the current density at the cross section with coordinate x:

$$j_x = \frac{j_0}{1+bx}, \tag{1}$$

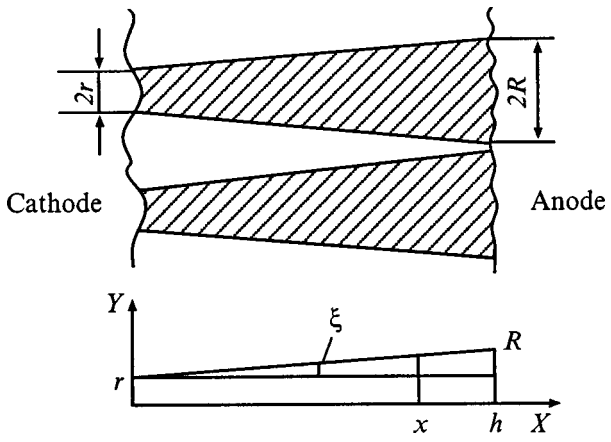

FIG. 2. Divergence diagram of the electron flow in the insulating gap in a plane parallel to the substrate surface.

where j_0 is the current density at the cathode surface and $b = \tan\xi/r$ is the variable parameter of the model.

We assume that neighboring electron beams do not overlap, i.e., the nanotips are sufficiently far apart on the cathode and the divergence angles are small; then, in the cross section at h there are no active (irradiated) regions with a current density different from Eq. (1) and the total current is obtained by summing the individual beams.

As in Ref. 7, we assume that the heat dissipated in the formed dielectric owing to the passage of the electron current through it is removed only to the substrate through a layer with a reduced thermal conductivity. Then the temperature T of the dielectric relative to the temperature T_0 of the substrate, which is regarded as a thermostat ($\Delta T = T - T_0$), at the cross section at x is given by

$$\Delta T_x = C j_x E_{\mathrm{fd}}, \tag{2}$$

where E_{fd} is the field strength in the formed dielectric and C is a constant.

The concentration ν_x (volume fraction) of particles in the conducting carbonaceous phase, which is determined by the local current density j_x and temperature ΔT_x (see below), also depends on the coordinate x: ν_x increases with distance from the cathode toward the anode. At some $x = h$ the concentration of particles in the conducting carbonaceous phase reaches a critical value ν_k equal to the percolation threshold. The corresponding values of ν_k for different spatial lattices lie within the range $0.3-0.6$.[8] This means that for $x > h$, the gap of the MIM structure contains a carbonaceous conducting medium which serves as an effective anode (Fig. 1). As scanning tunneling microscope measurements[4] of the potential in the insulating gap of a formed MIM diode show, this material has a very high conductivity and approaches graphite in terms of its characteristics.[9] Thus the voltage drop across the carbonaceous conducting medium can be neglected, and its boundary at $x = h$ becomes a moving anode.

The value ν_k corresponds to a critical temperature T_k (see below), in terms of which one can obtain from Eqs. (1) and (2) an expression for h, which determines the width of the insulating gap,

$$h = \frac{1}{b}\left[\frac{CjE_{\mathrm{fd}}}{(T_k - T_0)} - 1\right]. \tag{3}$$

Therefore, as opposed to Ref. 7, in the insulating gap of the MIM structure we now have three different segments in terms of their properties (Fig. 1): the dielectric of a tunnel barrier, the formed dielectric, and a carbonaceous conducting medium. The boundary between the first two is determined, as in Refs. 3 and 7, and that between the second and third, by percolation in the formed dielectric with a temperature and particle concentration in the conducting phase that depend on x. In principle, this variation may not only be related to the divergence of the electron flow, but can also be determined by a localized heat source (inside the formed dielectric), which leads to a drop in the temperature with distance from its center. However, the change in the flow density at the heat source is a more important factor; in addition, taking its finite size into account would require solving more compli-

cated heat transfer problems that are justified by our qualitative model. We shall return to this question later.

Another important approximation is the replacement of quantities characterizing the situation in the formed dielectric and dependent on the position x by constant averages over the formed dielectric region. This means that, having obtained an expression for the coordinate of the ''boundary of the carbonaceous conducting medium'' and thereby taking this qualitatively new factor into account, we again return to a uniform distribution of all quantities within the confines of the formed dielectric. This is allowable for small electron beam divergence angles, an assumption that appears to be justified for the high electric fields in the gap and also fits the qualitative features of the model.

In this approximation, the current density in the formed dielectric is given by

$$j_{\mathrm{fd}}=\frac{1}{2}(j_{\delta}+j_h), \tag{4}$$

and its temperature, by

$$T_{\mathrm{fd}}=CE_{\mathrm{fd}}j_{\mathrm{fd}}+T_0. \tag{5}$$

Here, by analogy with Ref. 7, we have

$$\nu_{\mathrm{fd}}=\frac{\alpha\sigma j_{\mathrm{fd}}}{ef\exp(-Q/\kappa T_{\mathrm{fd}})+\sigma j_{\mathrm{fd}}}, \tag{6}$$

where σ is the effective cross section for formation of a particle of the conducting carbonaceous medium from an organic molecule (dissociation) by electron impact, Q is the activation energy for the reaction in which this particle is lost, f is a constant with dimensions of frequency, e is the electron charge, and κ is Boltzmann's constant.

The critical temperature corresponding to the percolation threshold is found from an expression analogous to Eq. (6) for the current density and equals the temperature at the cross section at h,

$$T_k=\frac{Q}{\kappa\ln[ef/(\sigma j_h(\alpha/\nu_k-1))]}. \tag{7}$$

This last approximation can also be used, as before, to examine an insulating gap (Fig. 1) as a two-layer dielectric with constant relative permittivities (dielectric constants) of the layers, $\varepsilon_{\mathrm{fd}}$ and ε_b. Then the solution of the corresponding equations of electrostatics for the field in the formed dielectric gives

$$E_{\mathrm{fd}}=\frac{\Phi-\varepsilon\Phi_a+\varepsilon U}{\varepsilon h}, \tag{8}$$

where, as in Ref. 7,

$$\varepsilon=\frac{\varepsilon_{\mathrm{fd}}}{\varepsilon_b}=\frac{1}{(1-\nu_{\mathrm{fd}})^3}, \tag{9}$$

while the way to find the remaining quantities is clear from Fig. 1. It was assumed that the potential barrier at the cathode is triangular, i.e., imaging forces were neglected.

For the current density at the cathode we have used Simmons' formula,[10] which in our cases reduces to

$$j_0=\frac{A\Phi}{2\delta^2}\exp(-B\delta(\Phi/2)^{0.5}), \tag{10}$$

where A and B are constants and the width of the potential barrier, δ, at the Fermi level for the cathode is given by

$$\delta=\frac{\Phi}{\varepsilon E_{\mathrm{fd}}}. \tag{11}$$

As in Ref. 7, we have assumed that the current through the MIM structure is limited by tunneling at the cathode, so it agrees with Eq. (10) when the area of the emitting cathode surface is constant.

Equations (3) and (5)-(11) are a closed system of equations for the stationary state of the insulating gap of an MIM structure which we have solved numerically for the parameters h, T_{fd}, ν_{fd}, T_k, E_{fd}, ε, j_0, and δ as functions of U. The remaining parameters were varied near physically reasonable values.

DISCUSSION OF RESULTS

Some typical computational results from our model are shown in Fig. 3 (curves *1*). These curves were obtained for voltages U on the structure exceeding $\Phi_a=2$ V, since, as will be pointed out below, a smaller voltages there is a change in the mechanism for the conductivity of the insulating gap. As in Ref. 7, the current–voltage characteristic has a region with a negative differential resistance. The width of the insulating gap, which is now a variable quantity, decreases monotonically with the potential fall, and at the point $U=\Phi_a$ we have $h=\delta$, i.e., the formed dielectric region disappears (Fig. 1) and the boundary of the carbonaceous conducting medium (anode) rests at the edge of the cathode potential barrier. When the voltage is reduced further, the shape of the potential barrier will change from triangular to trapezoidal.

It should be noted that a change in the width h of the insulating gap represents the introduction of an additional feedback into the operating mechanism for nano-MIM diodes: yet another structural characteristic, h, which, along with the concentration of carbonaceous conducting particles and the barrier width, tracks changes in the voltage U and stabilizes the nanometer gap. In principle, this feedback can work independently of the first described in Ref. 7. In fact, let us consider a simplified model for nano-MIM diodes without any formed dielectric region at all. Here, in accordance with Eq. (10) the emission of electrons at the cathode will lead, by the same processes, to the formation of a carbonaceous conducting medium with a boundary determined by the same Eq. (3), but the insulating gap will consist of a uniform dielectric with a relative permittivity ε_b ($\varepsilon=1$), while the equations characterizing the formed dielectric vanish. It might be expected that for a given voltage on the structure, an equilibrium characterized by certain values of the gap width h and current density j_0 should develop, since h sets the electric field and the latter determines the barrier width δ and, thereby, the current density. The operation of such a device should involve simply a change in the gap

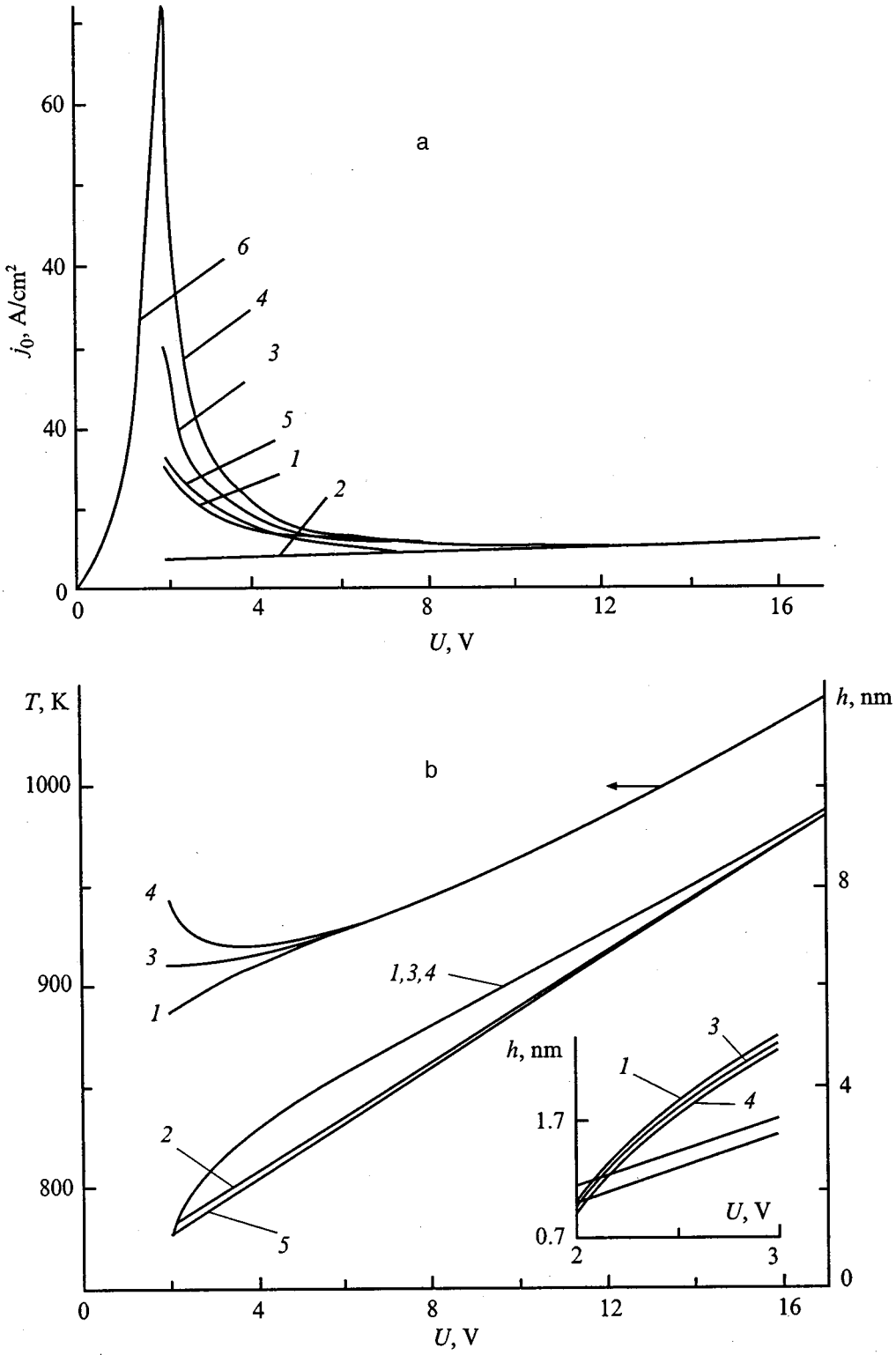

FIG. 3. Calculated dependences of the current density j_0, insulating gap width h, and temperature T of the formed dielectric: *1* — including percolation and the existence of the formed dielectric; $\Phi=2$ V, $\Phi_a=2$ V, $b=0.1$ nm^{-1}, $\nu_k=0.4$, $\alpha=0.6$, $\sigma=10^{-21}$ cm^2, $f=10^{13}$ s^{-1}, $C=10^{-8}$ cm$^3\cdot$deg$\cdot$W, $Q=2$ eV, and $T_0=300$ K; *2* — without a formed dielectric and with the same model parameters; *3* — with the formed dielectric and including the additional cooling effect, $l=3$ nm, $w=0.5$, and the remaining parameters the same; *4* — as curve *3*, but with $w=0.2$; *5* — without the formed dielectric, but including the additional cooling effect and with the same parameters as curve *4*; *6* — current–voltage characteristic of a tunnel gap with a constant width of $\delta=0.9$ nm and $\Phi=2$ V.

width as the voltage U varies. The corresponding system of equations (3), (7), (8), (10), and (11) yields stable solutions, which are shown in Fig. 3 (curve *2*) and confirm these ideas. This simplified mechanism retains the internal negative feedback, but, as can be seen from the graph, the current–voltage characteristic lacks a region with a negative differential resistance. This last difference is related to the loss of the modulation in the barrier width δ, which was caused by

stratification in the insulating gap, when the formed dielectric was eliminated.

This result is clearly illustrated by the following conclusion. In the general mechanism, we must isolate two different elements which are important for understanding the essence of the processes taking place in a formed nano-MIM diode with a carbonaceous active medium: internal negative feedback and modulation of the parameters (in this case, the width) of the potential barrier. The feedback loop, whose links include, in particular, the structural characteristics of the system (specifically the concentration of the particles of the carbonaceous conducting phase, the barrier width, and the width of the insulating gap), causes self-formation of the nanostructure in the carbonaceous medium as the nano-MIM diode operates. Since this feedback is negative (a current through the structure causes formation of particles of the conducting carbonaceous phase but a rise in the current leads to heating of the formed dielectric and the disappearance of the particles), the resulting nanostructure exists stably in an MIM-diode. These considerations led[6] to the concept of electroforming as a self-organizing process for a nanometer insulating gap and the gap itself was treated as a dissipative structure that develops in a carbonaceous conducting medium under certain conditions.

At the same time, the existence of feedback, as such, is not enough to create an N-shaped current voltage characteristic. That is ensured by modulation of the barrier width for tunneling electrons, i.e., by some mechanism for its ''forced'' variation, whereby the transparency of the barrier decreases as the voltage on the structure increases. In this case, this happens as a result of the stratification of the dielectric, but other mechanisms are possible. In particular, one such mechanism may be the increased transparency of the cathode potential barrier owing to imaging forces. In fact, as opposed to the situation described in Ref. 7, in the case of a moving anode, as the voltage U decreases the insulating gap is narrowed to very small widths (on the order of 1 nm for the model parameters used here). As h decreases, the imaging forces can cause a reduction in both the height and width of the barrier,[10] which represents a modulation in its transparency. Here we have neglected this effect, but it can evidently be important.

The mobility of the anode and the possibility of reducing the width of the formed dielectric region to zero in this model make the above approximation (whereby we neglect the finite size of the heat source in the formed dielectric, which has a finite thickness) too crude. There the heat flux through the side surface which bounds the carbonaceous conducting medium is neglected. For a formed dielectric of width $L=h-\delta$ comparable to its thickness l, which we can estimate as a few nanometers, the temperature ends up far too high. The resulting situation is clarified in Fig. 4. We assume that the heat source exists only in the formed dielectric region, since the mean free paths of the electrons in the disordered medium are very short and they lose all their energy within this region. Little heat is released in the carbonaceous conducting medium because of its high conductivity. We also assume that there is no heat flux through the vertical plane at the cross section at $x=\delta$; this is not well founded,

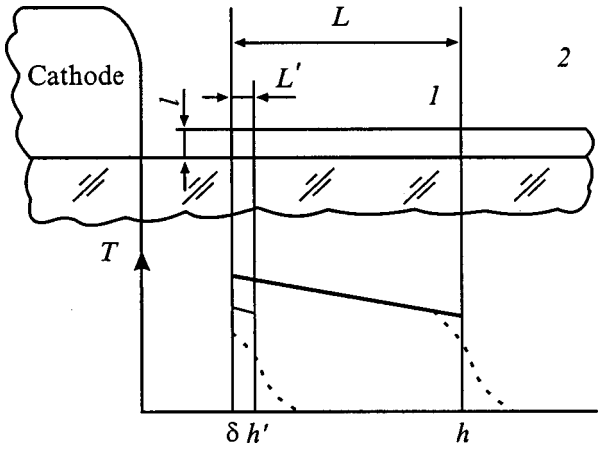

FIG. 4. A sketch of an insulating gap and the distribution of the temperature T in a formed dielectric *1* of thickness l (*2* — as in Fig. 1).

but it does simplify the discussion. The linear distribution of the temperature over x assumed in the above model is indicated in the figure by the solid lines. The actual distribution resulting from the finite size of the heat source and the loss of heat through the side surface is indicated by dashed curves; that is, at the boundary of the formed dielectric with the carbonaceous conducting medium the temperature distribution will be ''smeared out'' over a scale length determined by l. For large L this refinement is unimportant, but for $L=L'<l$ the situation changes. Besides the spatial smearing out of the temperature, its average value also decreases substantially in the formed dielectric, since the heat flux through the side surface becomes comparable to the main vertical flux into the substrate. This factor, which can be referred to as the additional cooling effect is most easily taken into account by introducing an effective heat transfer coefficient or, equivalently, replacing the constant C in Eq. (2) by an effective value C_{ef} defined as follows:

$$C_{\text{ef}}=\begin{cases} C & \text{for } L\geqslant l, \\ wC+\dfrac{L}{l}(1-w)C & \text{for } L<l. \end{cases} \tag{12}$$

Equation (12) simply represents a linear approximation for C_{ef} from C to wC as L is reduced from l to w, respectively. This approach allows us to retain the form of Eq. (2) over the entire range of variation of the parameters while taking qualitative account the additional cooling effect. Estimates show that, based on the atomic discreteness of the structure of the formed dielectric, the coefficient w can range between 1 and 0.1.

Calculations according to the complete model with Eq. (12) are shown in Fig. 3 (curves *3* and *4*). Their major feature is substantially higher current densities at low voltages U and, thereby, a more distinct region with a negative differential resistance. This seems natural, since considerably higher currents are required to maintain the critical temperature T_k because of the additional cooling effect. Less evident is the increase in the temperature of the formed dielectric (Fig. 3b) as the voltage U is lowered. This result becomes understandable, however, if we note that, according to Eq. (7), the criti-

cal temperature also increases as the current density increases. We will not present a plot of $\nu(U)$ here, but it is monotonic, decreasing from $\nu=\nu_k=0.4$ for $U=\Phi_a=2$ V to $\nu\approx 0.01$ for $U=15$ V. On the whole, it should be noted that including the additional cooling effect, with its obvious physical significance, introduces an independent parameter w which makes it possible to vary the steepness of the branches in the negative differential resistance over wide limits (cf. curves *1*, *3*, and *4*), something which was not possible up to now by varying any of the other model parameters. The resulting current–voltage characteristics are closer to experiment.[1]

The question of how the solution would look for a model with additional cooling but without a formed dielectric is of some interest. (That is, as above, we leave a uniform dielectric with a constant relative permittivity ε_b in the gap and we solve the contracted system of equations together with Eq. (12).) The results are illustrated in curves *5* of Fig. 3. It can be seen that, as opposed to curve *2*, the current–voltage characteristic has a region with a negative differential resistance, i.e., the additional cooling effect can be regarded as one of the mechanisms for modulating the barrier width for the tunneling electrons.

Finally, we return to voltages on the structure ranging from 0 to Φ_a. As noted above, here a cathode barrier will have a trapezoidal shape. The observed experimental maximum in the current–voltage characteristic is naturally related to the singular point $U=\Phi_a$. There, the barrier shape (Fig. 1) changes and the formed dielectric ($h=\delta$) vanish. It is important that the tunneling electrons no longer fall into the conduction band (or the impurity band) of the dielectric, but immediately show up in the conductivity band of the anode (carbonaceous conducting medium). Here the formation of conducting particles from organic molecules by electron impact becomes impossible, for the appropriate inelastic tunneling processes are evidently improbable. It is natural to suppose that as the voltage U is lowered, after passing the point $U=\Phi_a$, the width of the insulating gap ceases to vary. This is indicated indirectly by the experimental data of Ref. 4: in the highly conducting state of an MIM diode, even for a small voltage U, all the voltage falls, as before, across the nanometer insulating segment of the gap and is not spread uniformly over the gap length, as might be expected if the gap collapsed. But, then its current–voltage characteristic would simply have the standard form (10) only for low voltages (including the reverse tunneling process[10]). Figure 3a shows the corresponding calculations (curve *6*) for a gap width equal to the final value ($U=\Phi_a=2$ V) for curve *5* ($\delta=0.9$ nm). This branch is not shown for the remaining quantities simply because it is evident. The shape of curve *6* is in good agreement with experiment.[1] It should be noted that in this model, the maximum of the current–voltage characteristic corresponds to Φ_a, while its experimental value (3 − 4 V) is in satisfactory agreement with the values of Φ_a expected from physical considerations and which should be slightly lower than the work function for these metals.

We now present calculations in which the additional model parameters introduced in this paper (beyond those of Ref. 7) are varied. Varying b, which characterizes the divergence angle of the electron flux emitted from the cathode, has little effect on the current–voltage characteristic of the structure, but does significantly affect the $h(U)$ and $T(U)$ curves (Fig. 5). As the divergence angle increases, naturally, the temperature of the formed dielectric increases, since it is an average between T_k, which varies little for $U>4$ V, and the temperature at the cross section $x=\delta$. The width of the insulating gap then decreases, since the current density and temperature decrease more rapidly with x and the point at which T_k is reached comes closer to the cathode.

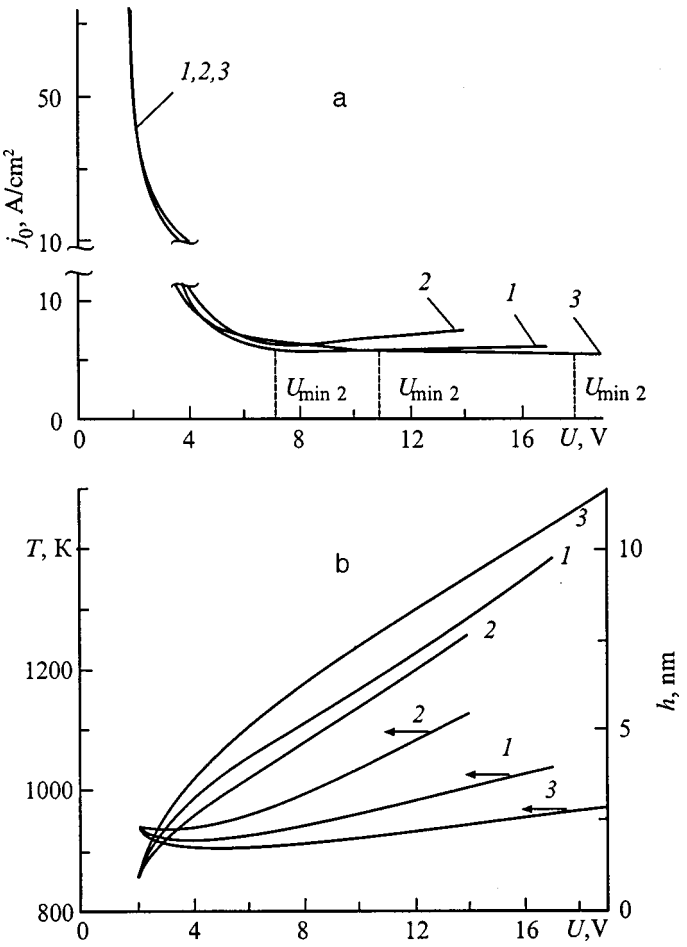

FIG. 5. Calculated variations in the current density j_0, insulating barrier width h, and formed dielectric temperature T. The conditions and parameters are the same as for curves *4* of Fig. 3; $b=0.1$ (*1*), 0.2 (*2*), 0.05 nm^{-1} (*3*).

Special attention should be paid to the behavior of the minimum in the current–voltage characteristic. The ascending branch of the characteristic in this region becomes much more mildly sloping than in Ref. 7, in better accord with experimental data.[1] This result is also a consequence of the variable gap width h: the field strength and, thereby, the current change very little here in comparison with the case of a fixed gap width. The presence of a minimum, as such, as that of the ascending branch of the current–voltage characteristic for high voltages, is related to the divergence of the electron beam. The slope of the curves in this region and the location of the minimum are controlled by the constant b, as is quite evident in Fig. 5.

Increasing the thickness l of the formed dielectric broad-

ens region where the temperature spreads out (Fig. 4) and produces a corresponding expansion of the range of voltages U in which the additional cooling effect operates. In the current–voltage characteristics this shows up as a reduction in the steepness of the branch with a negative differential resistance while retaining the maximum current; the dependences of the other quantities on U change little. Finally, increasing the percolation threshold ν_k produces a significant rise in the current maximum at $U=\Phi_a$, which steepens the characteristic significantly, in accordance with Ref. 7; here the $T(U)$ curve is shifted downward, parallel to itself, as a reflection of the drop in T_k. The associated graphs are not shown here because they provide little other information.

In evaluating the proposed model as a whole, we must note its advantages over the earlier model.[7] First, it agrees better with experiment, including the current–voltage characteristics of MIM diodes, and in terms of the whole series of elements mentioned above. Second, as it takes percolation into account, it provides a more correct description of the processes taking place in the insulating gap of an MIM structure. Thus, in particular, all the model parameters have physically reasonable values. Third, by varying these parameters, it is possible to vary the current–voltage characteristic over wide limits, which was difficult in the earlier model.[7] This meant that some important factor had not been taken into account. Despite the comparatively large number of selectable parameters, further development of the model and comparing it with more detailed experiments may provide independent and important estimates of these parameters at various operational limits for MIM diodes, each of which is described by only a few of these parameters. But even if this turns out to be impossible for some reason, the significance of the model, even in its present form, lies in the fact that it answers the question of which processes in MIM diodes can be responsible for their nontrivial features and in what manner.

Although the model is qualitative in character and uses rather crude approximations, its improvement through refinement of these approximations is quite inappropriate at this time. First, that would increase the number of parameters to be selected. Second, as the experience of the present calculations shows, the current–voltage characteristic is insensitive to improvements of this sort. This has a simple explanation. As noted previously,[7] the mechanism at the foundation of this model is based on a balance of the exponentials in Eqs. (6) and (10). Against the background of these very strong nonlinear dependences, most refinements will just yield some insignificant corrections.

[1] H. Pagnia and N. Sotnik, Phys. Status Solidi A **108**, 11 (1988).
[2] A. K. Ray and C. A. Hogarth, Int. J. Electron. **69**, 97 (1990).
[3] V. M. Mordvintsev and V. L. Levin, Zh. Tekh. Fiz. **64**(12), 88 (1994) [Tech. Phys. **39**, 1249 (1994)].
[4] H. Pagnia, N. Sotnik, and W. Wirth, Int. J. Electron. **69**, 25 (1990).
[5] V. M. Gaponenko, Izv. Vyssh. Uchebn. Zaved. Fiz. No. 5, pp. 115–120 (1992).
[6] K. A. Valiev, V. L. Levin, and V. M. Mordvintsev, Zh. Tekh. Fiz. **67**(11), 39 (1997) [Tech. Phys. **42**, 1275 (1997)].
[7] V. M. Mordvintsev and V. L. Levin, Zh. Tekh. Fiz. **66**(7), 83 (1996) [Tech. Phys. **41**, 679 (1996)].
[8] B. I. Shklovskiĭ and A. L. Efros, *Electronic Properties of Doped Semiconductors* [Springer-Verlag, New York (1984); Nauka, Moscow (1979), 416 pp.].
[9] H. Pagnia, Int. J. Electron. **69**, 33 (1990).
[10] J. G. Simmons, in *Tunneling Phenomena in Solids*, edited by E. Burstein and S. Lundqvist [Plenum Press, New York (1969), pp. 135–148; edited by V. I. Perel', Mir, Moscow (1973), pp. 131–142].

Translated by D. H. McNeill

Waveguide regimes of a graded-index planar waveguide with cladding

A. M. Shutyĭ, D. I. Sementsov, A. V. Kazakevich, and D. G. Sannikov

Ulyanovsk State University, 432700 Ulyanovsk, Russia
(Submitted June 23, 1998)
Zh. Tekh. Fiz. **69**, 74–79 (November 1999)

An exact dispersion relation for the guided modes of a four-layer planar structure with an exponential permittivity profile in the main waveguide layer (slab) is derived for the first time on the basis of the wave approach. Periodicity of the mode characteristics as functions of the cladding thickness is established. The dependence of the energy flux distributions of various modes in the waveguide structure on the thicknesses of the cladding and slab layers is investigated.
 [S1063-7842(99)01311-2]

INTRODUCTION

The operation of many planar waveguide structures relies on the application of a thin-film cladding on the surface of the main guiding layer (slab). For example, metallic claddings are used to create polarization filters, to separate TE and TM modes, and to enhance the localization of radiation in the slab. However, metallic claddings produce significant attenuation of the guided modes in their direction of propagation.[1] Dielectric claddings are widely used to correct the dispersion characteristics of the corresponding guided modes.[2] The theoretical analysis of guided-wave propagation of light in four-layer structures with a dielectric cladding of finite thickness has been aimed primarily at structures having a step distribution of the dielectric permittivity.[3–6] Waveguide regimes in four-layer graded-index waveguides with linear and parabolic permittivity profiles have been investigated[7,8] on the basis of approximate solutions of the electromagnetic field equations. An exponential permittivity profile is a close approximation to the profile formed in the thermal diffusion of metal ions on the surface of glass or optical crystals.[1,9] This profile can be used to approximate other profiles encountered in graded-index waveguides — a linear profile in particular. In this paper we consider the waveguide regimes of light transmission in a four-layer structure whose main guiding layer has an exponentially profiled permittivity. On the basis of an exact solution of the boundary-value problem and a numerical analysis of the resulting dispersion relation we investigate the specific characteristics of these regimes associated with the presence of a high-permittivity cladding layer.

FIELDS IN THE WAVEGUIDE STRUCTURE

The investigated planar waveguide structure consists of a slab waveguide with an exponential permittivity profile, a cladding layer, and the cover medium. The distribution of the permittivity along the normal to the surface of the slab is described by the equation

$$\varepsilon(x)=\varepsilon_3+(\varepsilon_1-\varepsilon_3)\exp(-x/a), \tag{1}$$

where ε_1 and ε_3 are the maximum and minimum permittivities of the slab, which occur at the interface with the cladding and in the bulk of the slab, respectively.

The region of the structure $x>a$ is regarded as the substrate; from now on we refer to the parameter a of the profile Eq. (1) as the thickness of the slab. The dielectric permittivity of the cladding is assumed to have the value ε_2, and the permittivity of the cover medium is ε_4; the following relation holds between the permittivities of the various layers: $\varepsilon_4\leqslant\varepsilon_3<\varepsilon_1<\varepsilon_2$. The structure is oriented so that the x axis is perpendicular to the interfaces, the slab–cladding interface occupies the plane $x=0$, and the interface between the cladding and the cover medium occupies the plane $x=-L$. The magnetic susceptibilities of all the layers are assumed to be equal to unity in the operating optical range.

We describe the field of a mode propagating in the waveguide along the z axis by the components

$$F_j(x,z)=C\mathcal{F}_j(x)\exp(-i\beta z), \quad j=x,y,z, \tag{2}$$

where C is a constant normalized to the radiation power transmitted into the waveguide, β is the propagation constant of the guided mode, and $\mathcal{F}_j(x)$ are profile functions, which characterize the field distribution of the mode along the thickness of the waveguide for the chosen distribution of the dielectric permittivity among the layers of the waveguide structure.

The wave equations describing the profile function in each of the media can be written in the form

$$\frac{d^2\mathcal{F}_j(x)}{dx^2}+[k_0^2\varepsilon(x)-\beta^2]\mathcal{F}_j(x)=0, \tag{3}$$

where $k_0=\omega/c$, ω is the frequency of the radiation, and c is the speed of light in vacuum; the permittivity $\varepsilon(x)$ is equal to ε_2 and ε_4 for the cladding and the cover medium, respectively, and is given by Eq. (1) for the slab and substrate (the region $x\geqslant 0$).

We now write the tangential components of the vector function $\mathcal{F}(x)$ relative to the interfaces. For a TE mode the component F_y has the meaning of the electric field component E_y, and for a TM mode it is the magnetic field component H_y. Here

$$\mathcal{F}_y(x)=\begin{cases} D\exp[q(x+L)], & x\leqslant -L, \\ A\cos hx+B\sin hx, & -L\leqslant x\leqslant 0, \\ J_{2w}[2v\exp(-x/2a)], & x\geqslant 0, \end{cases} \tag{4}$$

where we have introduced the notation $v=ak_0\sqrt{\varepsilon_1-\varepsilon_3}$, $w=ap$; $J_{2w}(\varphi)$ is the Bessel function of order $2w$ (Ref. 1).

The transverse components of the wave vector in the cladding, cover layer, and substrate are given by the expressions

$$h=\sqrt{k_0^2\varepsilon_2-\beta^2}, \quad q=\sqrt{\beta^2-k_0^2\varepsilon_4},$$
$$p=\sqrt{\beta^2-\varepsilon_3k_0^2}. \tag{5}$$

The z components of the profile functions for guided modes obey the relations

$$\mathcal{F}_z(x)=\pm\frac{i}{\gamma(x)k_0}\frac{d\mathcal{F}_y}{dx}, \tag{6}$$

where the upper sign corresponds to the H_z component of a TE mode, with $\gamma(x)=1$ in all of the media, while the lower sign corresponds to the E_z component of a TM mode, with $\gamma(x)=\varepsilon(x)$ in each of the media.

From the condition of continuity of the tangential components of the fields **E** and **H** at the interfaces we evaluate the constants in Eq. (4):

$$A=J_{2w}(2v), \quad B=-\frac{v\sigma}{ah}J'_{2w}(2v),$$
$$D=J_{2w}(2v)\cos(hL)+\frac{v\sigma}{ah}J'_{2w}(2v)\sin(hL). \tag{7}$$

Here we have $\sigma=1$ for a TE mode and $\sigma=\varepsilon_1/\varepsilon_2$ for a TM mode; the Bessel function derivatives used above are given by the expression $J'_{2w}(\varphi)=2wJ_{2w}(\varphi)/\varphi-J_{2w+1}(\varphi)$ (Ref. 10).

DISPERSION RELATION

The dispersion relation connecting the propagation constant of the corresponding guided mode to the parameters of the waveguide structure and the radiation can be written as follows in our case:

$$J_{2w}(2v)ah(q\delta-h\tan hL)+J'_{2w}(2v)v\sigma(q\delta\tan hL+h)=0, \tag{8}$$

where $\delta=1$ for TE modes, and $\delta=\varepsilon_4/\varepsilon_2$ for TM modes.

Without the cladding layer ($L=0$) Eq. (8) goes over to the well-known dispersion relation for a three-layer waveguide structure.[1,2] An analysis shows that the solutions of the above dispersion relation in the general case of a four-layer waveguide structure are governed by the set of mode orders $\nu=\nu_1+\nu_2$, where ν_1 corresponds to an integral number of energy minima of the mode field in the slab, and ν_2 corresponds to the same in the cladding. The mode indices ν_1 and ν_2 therefore determine the nature of the field distribution of a given mode order ν corresponding to the total number of energy minima in the waveguide structure.

The phase shift of a guided mode in the cladding is conveniently written in the form $Lh=(\nu_2+\alpha)\pi$, where to facilitate the analysis, we have introduced the dimensionless parameter α, which takes values from zero to unity. In this case the mode index ν_2 does not influence the form of the $\beta(a)$ curve, and the corresponding dispersion curves plotted for waveguide structures with identical parameters α and different values of ν_2 coincide. We introduce the thickness $L_h=\pi/h$, which is a function of the propagation constant β in general. In the range of the propagation constant $\sqrt{\varepsilon_3}k_0\leqslant\beta\leqslant\sqrt{\varepsilon_1}k_0$ the thickness L_h varies between the limits $\lambda/2\sqrt{\varepsilon_2-\varepsilon_3}\leqslant L_h\leqslant\lambda/2\sqrt{\varepsilon_2-\varepsilon_1}$, where λ is the light wavelength. If $(\varepsilon_1-\varepsilon_3)/\varepsilon_3\ll 1$ and if the permittivity of the cladding is much greater than ε_1 and ε_3, we can assume with a high degree of accuracy that L_h is constant. We then have $L=(\nu_2+\alpha)L_h$, where L_h is the characteristic period for the cladding thickness. For $\alpha=0$, i.e., for $L=\nu_2L_h$, the dispersion relation (8) goes over to the corresponding equation for a three-layer structure.

If the transverse component of the wave vector in the substrate is equal to zero ($p=0$, $\beta=k_0\sqrt{\varepsilon_3}$), the corresponding guided mode has a cutoff, i.e., it is transformed into a leaky mode. The effective cutoff thickness of the slab a_0 is determined from Eq. (8) and is given by the expression

$$a_0=\frac{J'_0(2v_0)v_0\sigma[q_0\delta\tan h_0L+h_0]}{J_0(2v_0)h_0[h_0\tan h_0L-q_0\delta]}, \tag{9}$$

where the transverse components $h_0=k_0\sqrt{\varepsilon_2-\varepsilon_3}$ and $q_0=k_0\sqrt{\varepsilon_3-\varepsilon_4}$; the parameter $v_0=a_0k_0\sqrt{\varepsilon_1-\varepsilon_3}$.

Inasmuch as v_0 is a function of the cutoff thickness, the determination of a_0 requires the numerical solution of Eq. (9). For modes with $\nu_1=0$, if $\alpha\geqslant\alpha_0$, the slab waveguide does not have a thickness a_0 corresponding to cutoff. It also follows from Eq. (8) that the interval $\alpha_{c1}\leqslant\alpha\leqslant\alpha_{c2}$ contains a thickness value a_0, above which the cladding becomes a waveguide layer for the mode, and its propagation constant lies in the interval $\sqrt{\varepsilon_1}k_0\leqslant\beta\leqslant\sqrt{\varepsilon_2}k_0$. The quantities a_c, α_0, and α_{c1} can be evaluated only by solving Eqs. (8) and (9) numerically.

NUMERICAL ANALYSIS

We now analyze the above-derived relations for the TE modes. We choose the parameters of the layers to match a real waveguide structure: A slab waveguide with the exponential profile (1) and a surface permittivity $\varepsilon_1=2.67$ is formed by the thermal diffusion of silver Ag_+ ions on a glass substrate with $\varepsilon_3=2.40$; for the cladding we choose a glassy chalcogenide semiconductor with a permittivity $\varepsilon_2=6.15$, and the cover medium is air with $\varepsilon_4=1$. The working wavelength of the laser beam is assumed to be $\lambda=0.633\,\mu$m; close to this wavelength we can assume that all the layers of the waveguide structure are nonabsorbing, and the permittivities of all the layers are real-valued.

Figure 1 shows the dependence of the propagation constant β on the slab thickness a for various TE mode orders and for various values of α: $\alpha=0$ (curves *5*,*7*,*9*), 0.15 (*4*), 0.2 (*3*,*6*,*8*), 0.23 (*2*), 0.25 (*1*). The calculations are carried out for mode indices $\nu_1=0,1,2$ (curves *1*–*5*, *6* and *7*, and *8* and *9*, respectively). For the investigated structure at the given wavelength the thickness L_h remains essentially con-

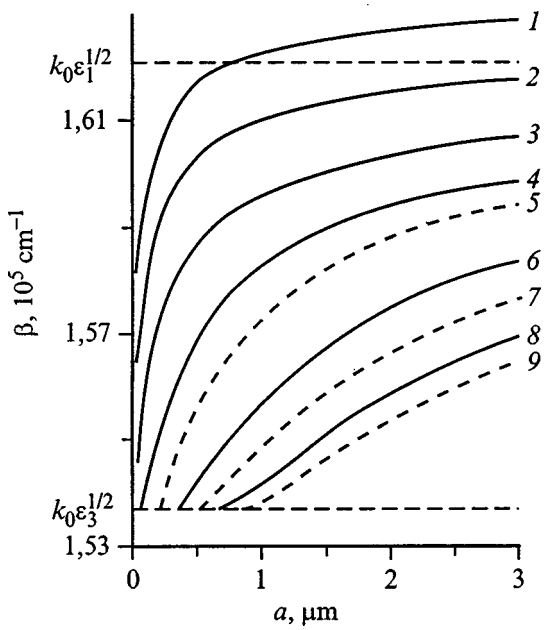

FIG. 1. Dependence of the propagation constant on the thickness of the slab waveguide.

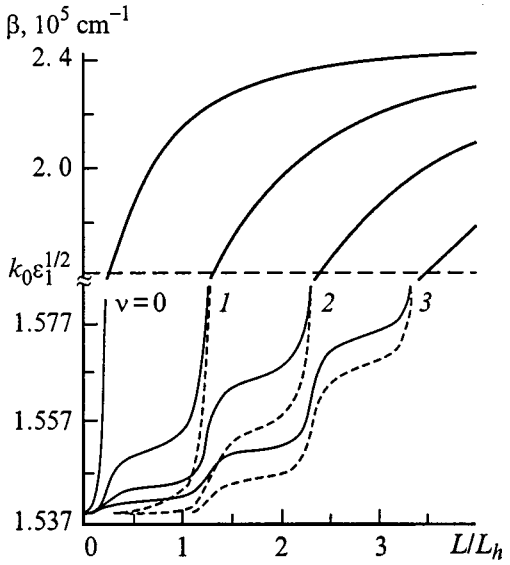

FIG. 2. Dependence of the dielectric permittivity for TE modes with $\nu=0-3$ on the normalized thickness of the cladding L/L_h.

stant and equal to $(0.157\pm0.002)\,\mu$m for all modes of the slab waveguide. It follows from the graphs that an added cladding with a higher permittivity than the slab reduces the mode cutoff thickness a_0 and increases the mode permittivities in comparison with a three-layer waveguide structure. For a cladding of thickness $L\cong\nu_2\cdot 0.157\,\mu$m, i.e., for $\alpha=0$, the dispersion curves for the modes of the four-layer waveguide structure coincide with those of the three-layer structure (dashed curves *5*, *7*, and *9*). In this case the mode with $\nu=\nu_2$ becomes the first mode ($\nu_1=0$). The lowest mode orders are absent from the given diagram, since they are cladding modes and become leaky modes for the slab. For modes with index $\nu_1=0$ a cutoff thickness a_0 does not exist for $\alpha>0.2$. The range of α in which slab modes are transformed into cladding modes is bounded by the endpoints $\alpha_{c1}=0.23$ and $\alpha_{c2}=0.302$.

Figure 2 shows the dependence of the permittivity for the first four modes with $\nu=0-3$ on the normalized cladding thickness L/L_h for various slab thicknesses a. The solid curves represent $a=0.05, 0.5, 0.84, 1.14\,\mu$m, which correspond to the cutoff thicknesses of the indicated modes in the unclad structure (for $L=0$), and the dashed curves represent mode orders $\nu=1-3$ for $a=0.3, 0.575, 0.9\,\mu$m. In accordance with the assumed representation of the thickness of the transverse layer $L=(\nu_2+\alpha)L_h$ the parameter α is equal to zero for an integer-valued L/L_h. It is evident from the curves that in the range of small α an increase in the cladding thickness leads to the most pronounced variation of the permittivity for TE modes. In the range $\beta\geqslant\sqrt{\varepsilon_1}k_0$ the guided modes become cladding modes, where now the layer with an exponential permittivity profile is the substrate of the waveguide structure.

Figure 3 shows the dependence of the cutoff thickness a_0 on the maximum value ε_1 (a) and the minimum value ε_3 (b) of the permittivity of the slab for TE mode orders $\nu=0-4$. Clearly, as ε_1 approaches the value $\varepsilon_3=2.40$, and as ε_3 approaches $\varepsilon_1=2.67$, we achieve the case of a weakly directional waveguide, which is characterized by an abrupt increase in the mode cutoff thickness.

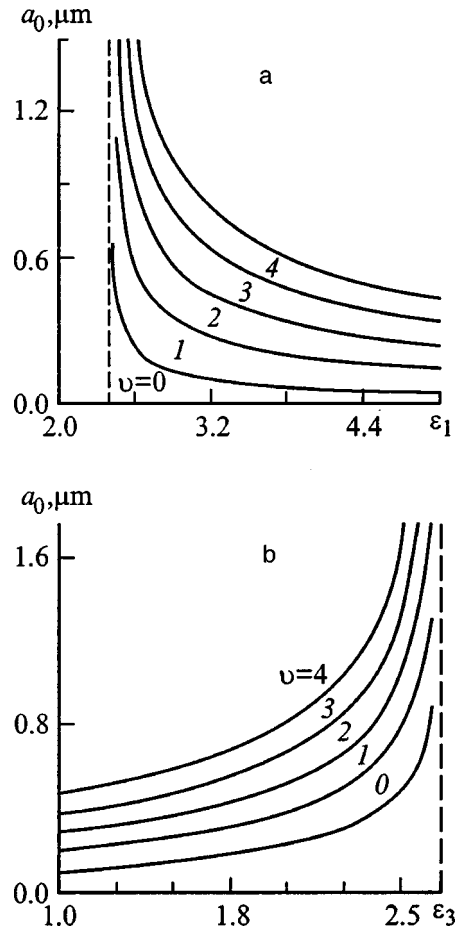

FIG. 3. Dependence of the cutoff thickness a_0 on the maximum value ε_1 (a) and the minimum value ε_3 (b) of the permittivity of the slab.

DISTRIBUTION OF ENERGY FLUXES

The distribution of the transported radiant energy flux over the cross section of the waveguide is given by the equation

$$\mathbf{S}(x)=\frac{c}{8\pi}\,\mathrm{Re}[\mathbf{E}(x)\times\mathbf{H}^{*}(x)]. \tag{10}$$

When the geometry of the investigated structure and the fields in the waveguide (2) are taken into account, Eq. (10) assumes the form

$$\mathbf{S}(x)=\mathbf{e}_z S_0\frac{\mathcal{F}_y^2(x)}{\gamma(x)}\left(\int_{-\infty}^{\infty}\frac{\mathcal{F}_y^2(x)}{\gamma(x)}dx\right)^{-1}, \tag{11}$$

where the unit vector $\mathbf{e}_z$ indicates the direction of the energy flux in the waveguide structure, and S_0 is the amplitude of the energy flux coupled into the waveguide (per unit length of the waveguide along the y axis).

If the layers are nonabsorbing, the energy flux does not have a component along the x axis.[11] Invoking the expressions for the longitudinal functions (4) and integrating $\mathbf{S}(x)$ over the coordinate x within the limits of the corresponding layer, we can find the linear energy flux density of the mode $\mathbf{S}_i$ in each of the four media. We introduce the localization factor

$$\eta_i=\frac{S_1}{S_0}=\left(\int_{L_i}\frac{\mathcal{F}_y^2(x)}{\gamma(x)}dx\right)\left(\int_{-\infty}^{\infty}\frac{\mathcal{F}_y^2(x)}{\gamma(x)}dx\right)^{-1}, \tag{12}$$

which characterizes the fraction of the energy flux transported by the mode in the ith layer. The effective thickness of the waveguide for slab modes depends on the sum of the thickness of the slab, the thickness of the cladding layer, and the small penetration depth into the cover medium, which can be disregarded in most cases.

Figure 4 shows the distribution of the energy flux $S(x)$ transported by the mode $\nu=0$ in a structure with a slab of thickness $a=3\ \mu$m for various cladding thicknesses L. The amplitude of the mode field is chosen such that the linear energy flux density propagating along the waveguide is equal to 1 erg/s·cm. For the construction of these curves we have chosen cladding thicknesses $L=0.157\alpha\ \mu$m, which correspond to values of the parameter $\alpha=0, 0.15, 0.2, 0.23$ (Fig. 4a, curves *1–4*) and $0.25, 0.3, 0.5$ (Fig. 4b, curves *5–7*). Curves *1–4* refer to slab modes, and curves *5–7* represent cladding modes. The dashed lines at $x=-L/a$ correspond to the interface of the cladding with the cover medium. It follows from the graphs that increasing the cladding thickness shifts the maximum of the energy flux closer to the slab–cladding interface. As a result, strong localization of the mode energy near the interface can be achieved for modes of the investigated type ($\nu=0$) at a thickness close to a_c ($a_c=3\ \mu$m for $\alpha=0.236$). A further increase in the thickness L converts the slab modes into cladding modes, where the energy maximum continues to increase, the mode energy is localized mainly in the cladding, and the thickness of the guiding structure diminishes drastically, approaching the thickness of the cladding layer.

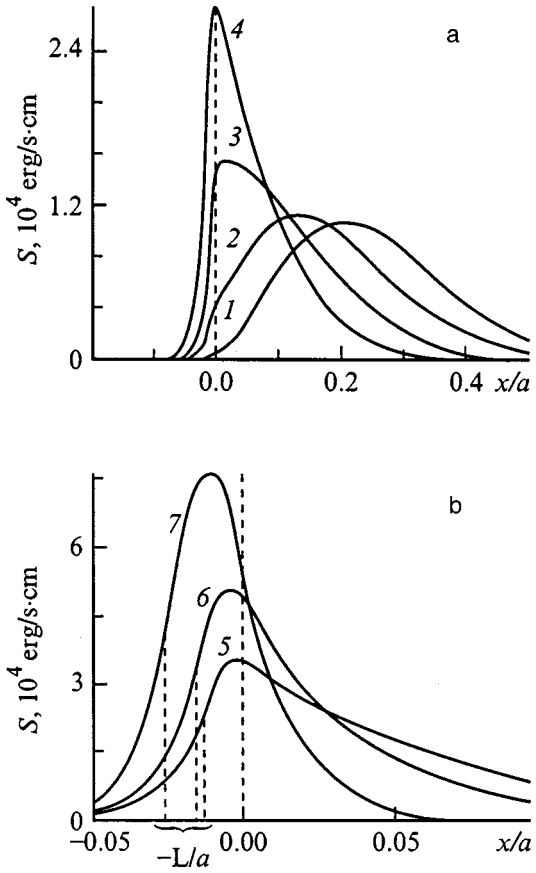

FIG. 4. Distribution of the energy flux transported by the fundamental ($\nu=0$) TE mode over the cross section of the waveguide structure.

The energy flux distribution for slab modes with $\nu=1$ is shown in Fig. 5. The corresponding curves are plotted for the parameters $\nu_1=1$, $\nu_2=0$ (a) and $\nu_1=0$, $\nu_2=1$ (b); $\alpha=0$ (curves *1,4*), 0.2 (*2*), 0.22 (*5*), 0.27 (*6*), 0.3 (*3*). It is evident that the corresponding variation of the parameter α as the cladding thickness L is increased causes the energy minimum in the slab to shift toward the interface with the cladding, and in the limit $\alpha=1$ the slab mode with indices ν_1 and ν_2 becomes a mode with new indices ν_1-1 and ν_2+1. The total number of energy minima in the waveguide structure, which is dictated by the mode order ν, is constant in this case. The dependence of the fraction of the TE mode energy localized in the cladding on the parameter α is most pronounced in the interval $\alpha\sim 0-0.35$, where the influence of the cladding on the mode propagation constants is the strongest. The maximum fraction of the mode energy in the cladding corresponds to $\alpha\approx 0.23-0.3$, for which the energy maximum is situated at the slab–cladding interface; the fraction of energy in the cladding is a minimum for $\alpha\approx 0.7-0.8$, when the energy minimum occurs at the interface. Close to the mode cutoff thickness a_0 the presence of the cladding can significantly enhance the degree of mode localization, producing severalfold variations in the values of the energy maxima. Here the influence of the cladding on the energy distribution of weakly localized modes in the waveguide structure is strong even for $\alpha>0.7$.

Adding one to the index ν_2 for constant values of α and

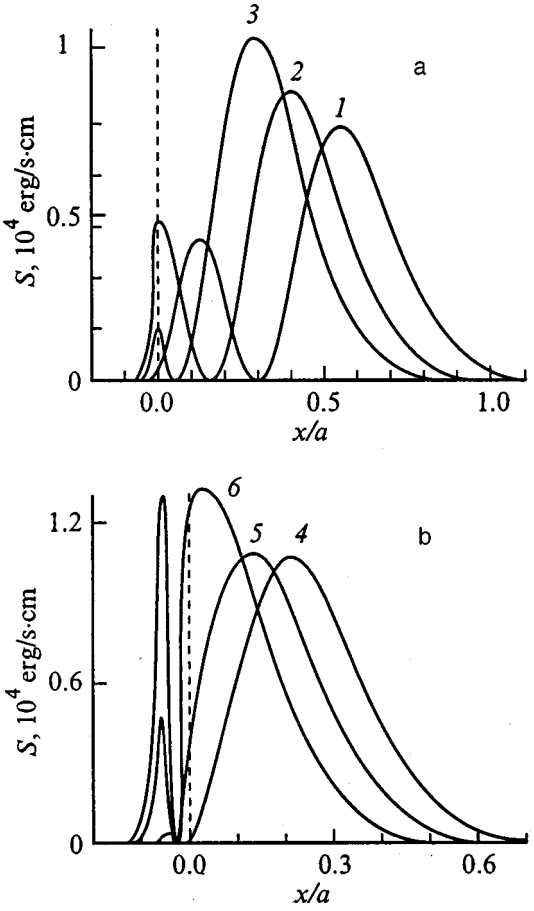

FIG. 5. Distribution of the energy flux of TE modes of the slab with indices $\nu_1=1$, $\nu_2=0$ (a) and $\nu_1=0$, $\nu_2=1$ (b) in a structure with $a=3$ μm.

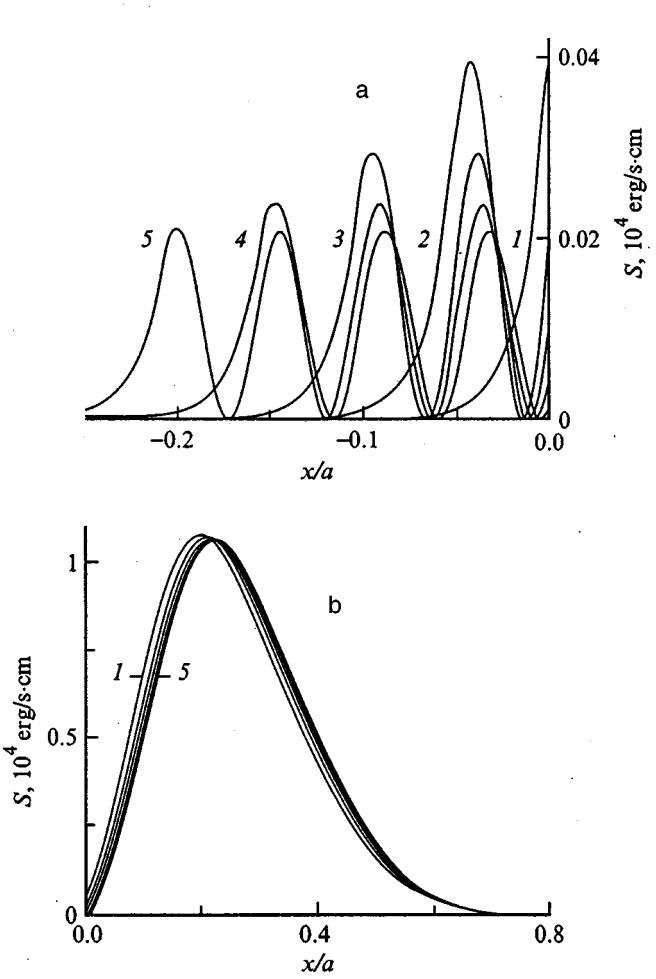

FIG. 6. Distribution of the energy flux in the cladding (a) and in the slab and the substrate (b).

ν_1 produces one more energy maximum (minimum) in the cladding. The distribution of the mode energy flux in the slab and the substrate remains essentially constant in this case. For $L=\nu_2 L_h$, i.e., when $\alpha=0$, the flux distribution coincides with the flux distribution of the mode with $\nu=\nu_1$ in a three-layer waveguide structure as long as the reduction of the energy fraction in these layers is disregarded. Figure 6 shows the distribution of the energy flux in the cladding (a) and in the slab and the substrate (b) for $\alpha=0$, $\nu_1=0$, and $\nu_2=0-4$ (curves *1–5*).

CONCLUSION

The foregoing analysis shows that the presence of an additional cladding layer with a high dielectric permittivity in a waveguide structure significantly influences all the mode characteristics of the waveguide. Their strong dependence on the thickness of the added layer can be utilized to effectively control the energy flux distribution of the corresponding mode over the cross section of the waveguide structure, maximizing the localization of the mode for the given permittivity of the slab waveguide. We also note that the real permittivity profile of asymmetric waveguide structures prepared by the thermal diffusion of metal ions into a glass substrate is well approximated by an exponential profile.

[1] M. Adams, *An Introduction to Optical Waveguides* [Wiley, New York (1982); Mir, Moscow (1984), 512 pp.].
[2] *Guided-Wave Optoelectronics*, edited by T. Tamir (Springer-Verlag, Berlin–New York, (1988); Mir, Moscow (1991), 575 pp.].
[3] R. F. Carson and T. E. Batchman, Appl. Opt. **29**, 2769 (1990).
[4] P. V. Adamson, Opt. Spektrosk. **70**, 211 (1991) [Opt. Spectrosc. **70**, 121 (1991)].
[5] H. Ding, P. Gerard, and P. Benech, IEEE J. Quantum Electron. **31**, 411 (1995).
[6] P. A. Solana and Yu. P. Udoev, Zh. Tekh. Fiz. **61**(8), 198 (1991) [Sov. Phys. Tech. Phys. **36**, 960 (1991)].
[7] She Shouxian, Wang Jindyi, and Qiao Li, Opt. Commun. **90**, 238 (1992).
[8] V. K. Chaubey, K. K. Dey, and P. Khastgir, J. Opt. Commun. **15**, 95 (1994).
[9] V. N. Chebotin, *Chemical Diffusion in Solids* [in Russian], Nauka, Moscow (1989), 207 pp.
[10] E. Jahnke and F. Emde, *Tables of Higher Functions*, 6th ed. rev. by F. Lösch [McGraw-Hill, New York (1960); Nauka, Moscow (1977), 342 pp.].
[11] D. I. Sementsov and A. M. Shutyĭ, Opt. Spectrosc. **81**(1), 153 (1996). [Opt. Spectrosc. (Russia) **81**, 139 (1996)].

Translated by James S. Wood

Sensor based on a $Pt/LaF_3/SiO_2/SiC$ structure for the detection of chlorofluorocarbons

V. I. Filippov, A. A. Vasil'ev, and A. A. Terent'ev

Kurchatov Institute Russian Science Center, 123182 Moscow, Russia

W. Moritz and U. Roth

Humboldt-Universität, 10117 Berlin, Federated Republic of Germany

(Submitted July 22, 1998)
Zh. Tekh. Fiz. **69**, 80–85 (November 1999)

A metal–insulator–semiconductor structure based on silicon carbide with a subgate layer of LaF_3 solid electrolyte is discussed as a gas sensor. The kinetics of the variation of the flat-band potential of a $Pt/LaF_3/SiO_2/SiC$ structure in interaction with chlorofluorocarbons (Freons) is investigated in the temperature range from 300 to 530 °C. The activation energies of the gas sensitivity are estimated from the temperature dependences of the response rate of the sensor to various Freons. The possibility of detecting all the investigated chlorofluorocarbons at a concentration level of 10 ppm in air is demonstrated. © *1999 American Institute of Physics.* [S1063-7842(99)01411-7]

INTRODUCTION

The presence of chlorofluorocarbons (Freons) in the atmosphere is highly destructive to the earth's ozone layer. Despite the ban against compounds of this kind, Freons continue to enter the atmosphere, inflicting significant damage upon the ecology of the planet. Chemical sensors with a sensitivity to fluorine-containing gases hold considerable promise for monitoring the indicated processes when it is required to know the concentration of pollutants in the air. Structures fabricated from metal-oxide semiconductors of the type $V–Mo–Al_2O_3/ZnO$ or stannic oxide are sensitive to such Freons as CF_3CH_2F and CCl_2FCClF_2 (Ref. 1). However, sensors that utilize metal oxides react not only to the presence of Freons in the air, but to other gases as well, hydrocarbons in particular, i.e., they are nonselective in their sensitivity to chlorofluorocarbons. The selectivity can be enhanced slightly by doping the gas-sensitive layer with sulfur.[2] It should also be noted that for sensors of this kind to exhibit stable operation, they require the presence of oxygen in the investigated gaseous medium. This consideration can severely limit the domain of application of gas sensors based on metal-oxide semiconductors.

We have previously[3] described a sensor that can be used to record the fluorine and hydrogen fluoride content both in free air and in an inert medium; its operating principle is based on the field effect. The active components of the $Pt/LaF_3/Si_3N_4/SiO_2/Si$ structure were a platinum electrode and a layer of LaF_3 solid electrolyte, which together with the investigated gas formed a three-phase boundary. It has been shown that a sensor based on such a structure is sensitive at a temperature of 180 °C to fluorine-containing gases, in particular to 1.1,1.2-tetrafluoroethane. However, this working temperature has been found to be insufficient for stable operation of the instrument in the detection of other chlorofluorocarbons. Unfortunately, the highest possible working temperature of sensor structures utilizing silicon, which has a 1.15-eV band gap, does not exceed 200 °C. By replacing the silicon with wide-gap silicon carbide (which has a 2.9-eV band gap) the working temperature of a $Pt/LaF_3/SiO_2/SiC$ structure can be raised to 530 °C. In the interval from room temperature to 200 °C the sensor signal is weakly dependent on the working temperature, and only a slight change is observed in the sensor response kinetics. An estimate of the effective activation energy from the initial rate of the sensor response process after exposure of the structure to fluorine gives a value of 0.2 eV.

The objective of the present study is to investigate the influence of temperature on the characteristics of a $Pt/LaF_3/SiO_2/SiC$ sensor in the measurement of Freon concentrations. We consider such characteristics as the sensitivity, the initial rate of change of the bias voltage of the structure when exposed at time t_0 to a concentration ''step'' of the investigated gas $\partial \Delta U/\partial t|_{t=t_0}$, and, finally, the selectivity of the sensor.

METHODS OF INVESTIGATION, DEVICES, AND EQUIPMENT

The structure of the gas sensor is described in detail in Ref. 4. A layer of solid LaF_3 (of thickness 240 nm) is deposited by thermal evaporation onto a silicon carbide semiconductor substrate[1)] with a 6*H*-SiC epitaxial layer (*n*-type, carrier density $10^{15}-10^{16}\text{cm}^{-3}$, thickness 5 μm) and a SiO_2 layer (35 nm). A catalytically active platinum electrode (of diameter 0.7 mm and thickness 30 nm) is then deposited by magnetron sputtering. A high-temperature Ohmic contact of burned-in nickel is formed on the opposite side of the substrate.

The $Pt/LaF_3/SiO_2/SiC$ gas-sensing structure is shown schematically in Fig. 1a, the distribution of the potential $\varphi(x)$ along the thickness of the structure is shown in Fig. 1b, and its capacitance–voltage curve is shown in Fig. 1c. As a result of processes taking place at the three-phase boundary (between the investigated gas, the catalytically active metal,

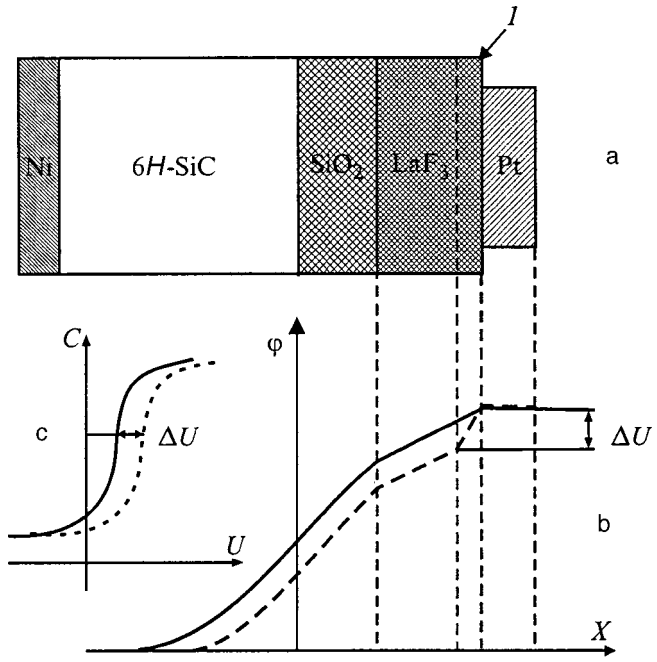

FIG. 1. Schematic view of the sensor structure (a), showing the charge double layer (*1*); distribution of the electronic potential along the structure (b); high-frequency capacitance–voltage curve of the structure (c) in the absence of fluorine (solid curves) and when exposed to fluorine (dashed curves).

and the solid electrolyte), a charge double layer is formed in the subsurface region of the LaF_3. This alters the potential distribution and, as a consequence, shifts the capacitance–voltage curve. If a simple oxidation–reduction reaction takes place at the three-phase boundary, the difference ΔE between the steady-state values of $\Delta U_{1,2}$ at two different concentrations of the investigated gas is related to these concentrations by the Nernst equation

$$\Delta E = \Delta U_2 - \Delta U_1 = (kT/Ze) \cdot \ln C_2/C_1, \quad (1)$$

where k is Boltzmann's constant, Z is the number of electrons taking part in the reaction, e is the electron charge, and C_1 and C_2 are the initial and final concentrations of the gas.

The quantity ΔU can be determined from the shift of the high-frequency $C-V$ curve of the structure. The method used in our work is to measure the variation of the bias voltage of the structure at a fixed capacitance. To maintain a constant capacitance during exposure of the structure to the gas in this case, the bias voltage U corresponding to the given capacitance is incremented by the amount ΔU, which is recorded by the instrumentation. Information about the characteristics of the gas-sensing structure is obtained by processing the relaxation curves $\Delta = f(t)$ recorded by the above-described procedure after the input of a concentration step of various Freons in air into the structure in the temperature interval from 300 to 530 °C. Figure 2 shows a block diagram of the experimental arrangement. A gas chamber containing the structure *1* is mounted in an external heater *2*. The working temperature in the gas chamber is monitored by

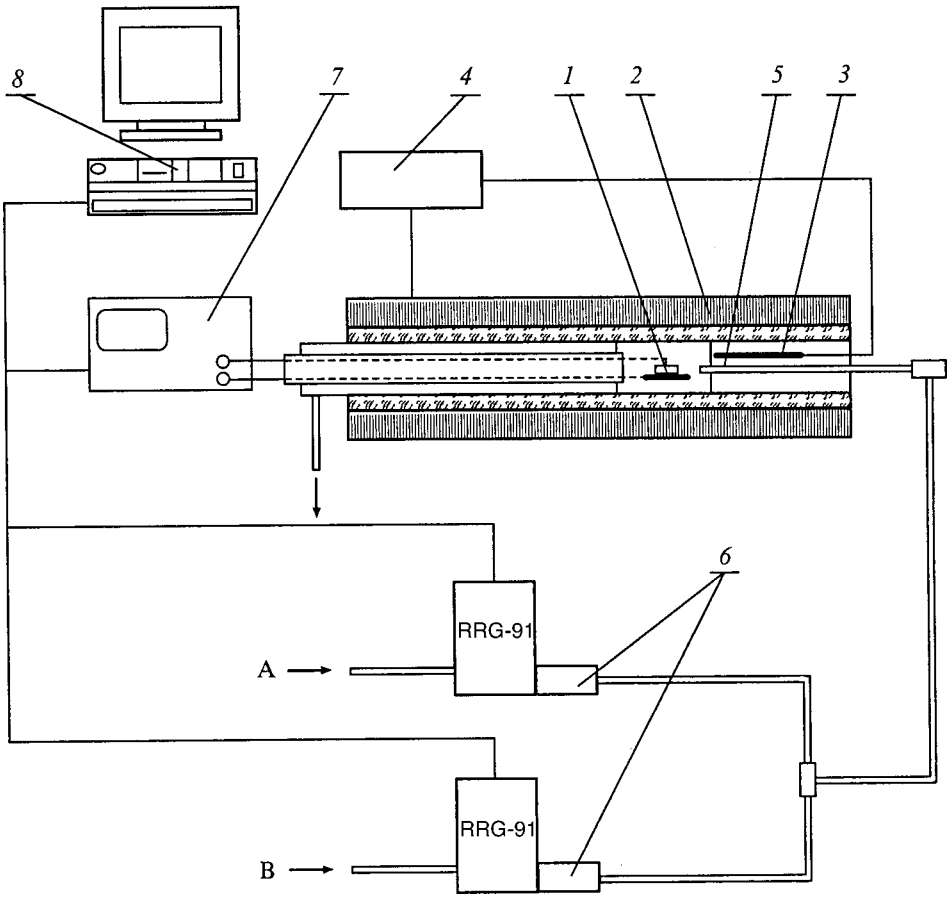

FIG. 2. Block diagram of the experimental apparatus: (A) fluorocarbon; (B) synthetic air.

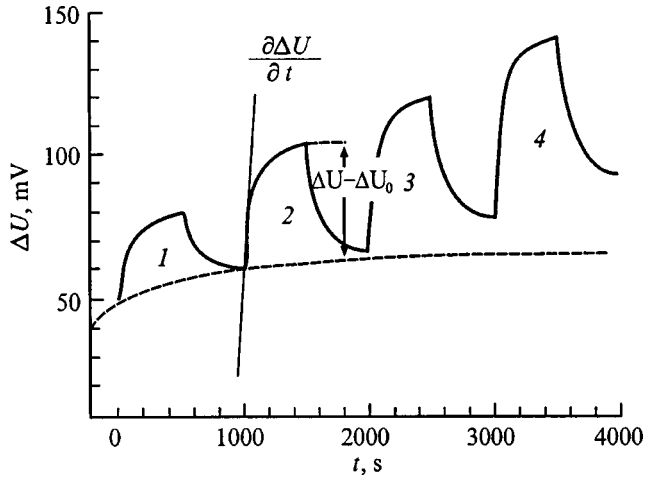

FIG. 3. Relaxation curve $\Delta U = f(t)$ after the input of a Freon concentration step into the gas chamber. F-11 concentration in air: *1* — 1000 ppm; *2* — 2000 ppm; *3* — 3000 ppm; *4* — 5000 ppm. Baseline concentration $C_0 = 500$ ppm, pulse duration 500 s, temperature $T = 458\,^{\circ}\mathrm{C}$.

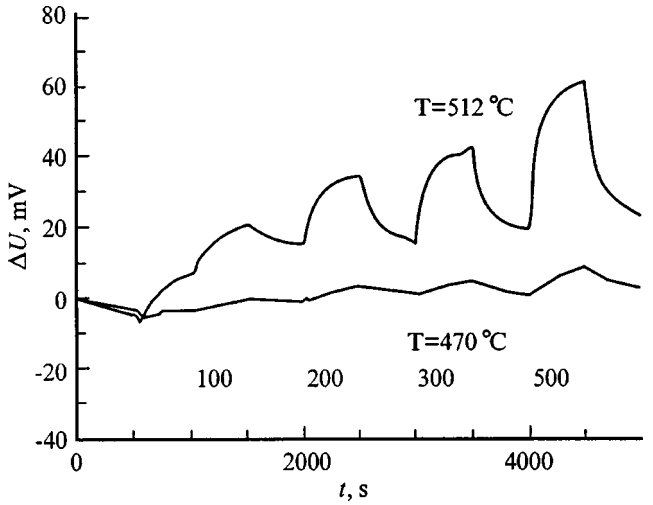

FIG. 4. Relaxation curves in response to 100 ppm, 200 ppm, 300 ppm, and 500 ppm F-11 in air. Baseline concentration $C_0 = 500$ ppm.

a Pt–Pt + 10%Rh thermocouple *3* and is controlled by the electronics section *4*. A gas mixture prepared by dynamic mixing of synthetic air and Freon is injected into the chamber by way of a capillary tube *5*. The required concentration of the investigated gas is set by RRG-91 electronic flow-rate regulators *6*. The dimensions of the capillary tube (ϕ 0.8 mm, $L = 25$ cm) are such as to maintain the temperature of the gas at the input to the chamber equal to the temperature of the heater at a gas flow rate of 2 cm^3/s. The materials of the chamber and the capillary tube are chosen to prevent catalytic dissociation of the Freons at the walls. The sensor signal, i.e., the variation of the bias voltage on the structure at a fixed capacitance, is recorded by an HP 4284A *LCR* meter *7* and is processed by an IBM PC 486 computer (*8*, which controls the entire operation of the apparatus. Special software has been developed to control the variation of the Freon concentrations and to generate the relaxation curves in automatic mode. In this study we have investigated air mixtures of the chlorofluorocarbons CCl_3F (F-11), $CHClF_2$ (F-22), CF_3CCl_3 (F-113), CF_3CH_2Cl (F-133), and CF_3CH_2F (F-134) and also tetrachloromethane and methane.

RESULTS AND DISCUSSION

Figure 3 shows a characteristic relaxation curve ($T = 458\,^{\circ}\mathrm{C}$) for the input of concentration steps of Freon in air (1000, 2000, 3000, and 5000 ppm) into the gas chamber at a baseline concentration $C_0 = 500$ ppm. The same figure also shows how the sensor characteristics are determined: The sensitivity ΔE (from the difference between the steady-state values of the change in the bias voltage) and the initial rate of change of the bias voltage $\partial \Delta U/\partial t|_{t=t_0}$ (from the derivative on the initial part of the relaxation curve). Analogous curves have been obtained for all the investigated Freons in the working temperature range. An increase in the temperature has the effect of increasing the sensitivity and the response rate of the sensor, as is evident in the example of F-11 (Fig. 4).

Figure 5 shows the sensitivities of the sensor to F-11 (458 °C), F-113, and F-133 (450 °C) as functions of the concentration C of the investigated gas at a baseline concentration of 500 ppm. It is seen that the sensitivity depends linearly on the logarithm of the concentration. This behavior is consistent with the Nernst equation (1), but the formal number of electrons involved in the reaction varies from 2 to 5 for different gases and can vary with the temperature (see Table I). This result indicates that the processes taking place at the three-phase boundary under the given experimental conditions have a different character than in fluorine detection,[3] where $Z = 2$.

Figure 6 shows the temperature dependence of the initial rate of change of the bias voltage for two chlorofluorocarbons. On the assumption that the temperature dependence has an activational character,

$$\left.\frac{\partial \Delta U}{\partial t}\right|_{t=t_0} \approx \mathrm{const}_1 \cdot \exp\left(\frac{-\Delta\varepsilon}{kT}\right), \tag{2}$$

these curves are plotted in Arrhenius coordinates. The activation energy is estimated to be $\Delta\varepsilon = 1 \pm 0.2$ eV. For the

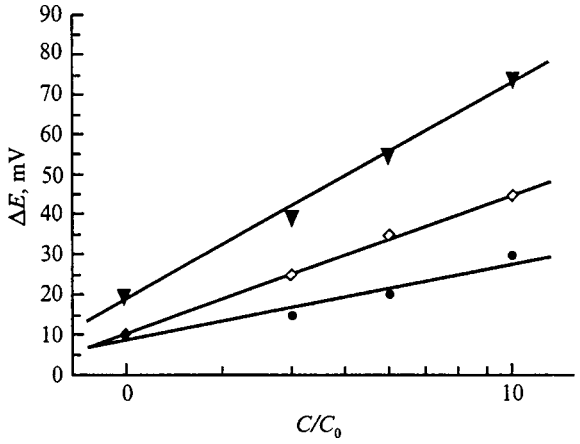

FIG. 5. Dependence of the sensor sensitivity ΔE on the concentration of F-11 (▼), F-113 (◇), and F-133 (●). Baseline concentration $C_0 = 500$ ppm.

TABLE I.

Freon	Temperature, °C	ΔE, mV/dec.	Z	α	$\Delta\varepsilon$, eV
F-11 (CCl_3F)	458	79	2	0.51	1.0
F-134 (CF_3CH_2F)	450	62	2	0.79	1.6
F-22 ($CHClF_2$)	458	29	5	0.64	0.8
F-133 (CF_3CH_2Cl)	450	28	5	0.75	1.3
F-113 (CF_3CCl_3)	450	51	3	1.2	1.0
	411	43	3	1.1	
	390	55	2	1.1	
	340	60	2	0.9	

majority of the gases (see Table I) the activation energy is much higher than its counterpart for fluorine. Moreover, in order of magnitude the activation energy is close to the activation energy of the pyrolysis reaction of complex molecules.[5] We can assume from this fact that the sensitivity of the investigated structure to Freons involves the preliminary dissociation of molecules at the three-phase boundary and the subsequent detection of dissociation products.

It has been shown previously[4] that the derivative at the initial time can be correlated with the concentration of the investigated gas C and can serve as a sensor signal. From the standpoint of practical applications of the sensor for the measurement of Freon concentrations it is important to analyze the concentration dependence of the derivative $\partial\Delta U/\partial t|_{t=t_0}$. We can now write Eq. (2) in the form

$$\left.\frac{\partial\Delta U}{\partial t}\right|_{t=t_0} \approx \mathrm{const}_2\cdot\exp\left(\frac{-\Delta\varepsilon}{kT}\right)\cdot(C-C_0)^{\alpha}. \tag{3}$$

Figure 7 shows the concentration dependence of the initial rate of change of the bias voltage in log–log scale for Freon 113 at several different temperatures. It is seen from estimates of the coefficient α that this dependence is almost linear for F-113. The results for the other gases are summarized in Table I.

An estimate of the lowest detectable Freon concentrations demonstrates the feasibility of stable detection of changes in their concentration in air at the level of 10 ppm at a temperature of 500 °C. However, at zero baseline concentration, i.e., when synthetic air is initially injected into the gas chamber, a signal of opposite polarity has been observed for certain samples with Freon concentrations at the level of several parts per million. It is conceivable that what is detected in this case is a change in the oxygen concentration (or the degree of oxidation of LaF_3 at the three-phase boundary) is under the influence of a small quantity of dissociation products of Freon molecules formed at the interphase boundary. Prevention of this effect requires stabilization of the properties of the subsurface layer of the solid electrolyte, for example, by pretreatment of the structure with fluorine.

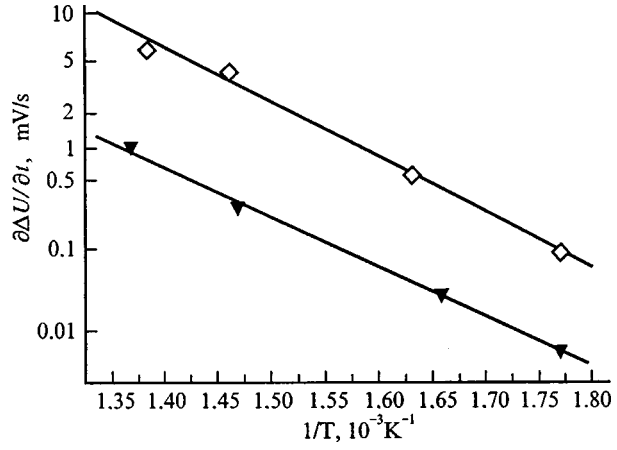

FIG. 6. Initial rate of change of the voltage on the structure versus reciprocal temperature after the input of F-11 (▼) and F-113 (◇) into the gas chamber.

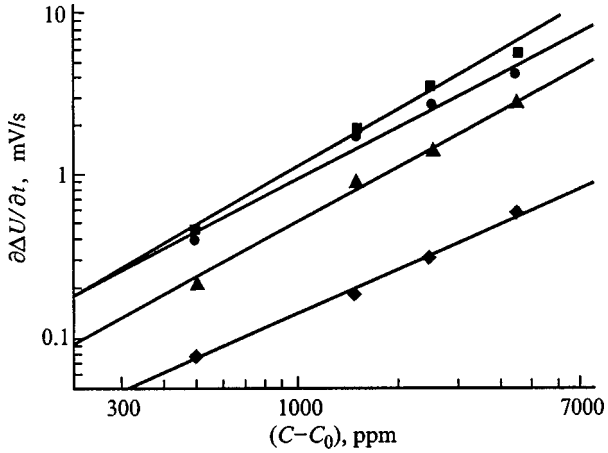

FIG. 7. Initial rate of change of the voltage on the structure versus F-113 concentration at $T=450$ °C ($\alpha=1.2$) (■), 411 °C ($\alpha=1.1$) (●), 390 °C ($\alpha=1.1$) (▲), and 340 °C ($\alpha=0.9$) (◆). The power exponent α in Eq. (3) is estimated.

Tetrachloromethane and methane were used in the experiment to investigate the selectivity of the sensor structure. Figure 8 shows the response of the sensor to the input of a concentration step of 500 ppm CCl_4 in air at two temperatures. An increase in the temperature leads to a substantial change in the response kinetics but has essentially no influence on the steady-state signal of the sensor. The nature of the sensitivity of the structure to tetrachloromethane is still not very well understood. This molecules does not have a dipole moment, and so the space-charge region in the semiconductor is not affected by molecules adsorbed at the Pt/LaF_3 interface.[6] Also, it would be difficult to expect any direct sensitivity of the structure to chlorine. Nonetheless, to explain the nature of the sensitivity of the structure to CCl_4, it would be most interesting to conduct experiments with chlorine; unfortunately, such an undertaking exceeded the experimental capabilities of the present study.

Figure 9 shows relaxation curves of the increment in the bias voltage upon the input to the structure of a step of 500 ppm of methane in air. At both temperatures the polarity of the signal is opposite to the so-called fluorine response. The nature of the response signal, specifically a shift of the bias voltage toward negative values, is similar to the response of an ordinary metal–insulator–semiconductor (MIS) structure (without the LaF_3 subgate layer) to hydrogen.[7] Sensors utilizing MIS structures with catalytically active elec-

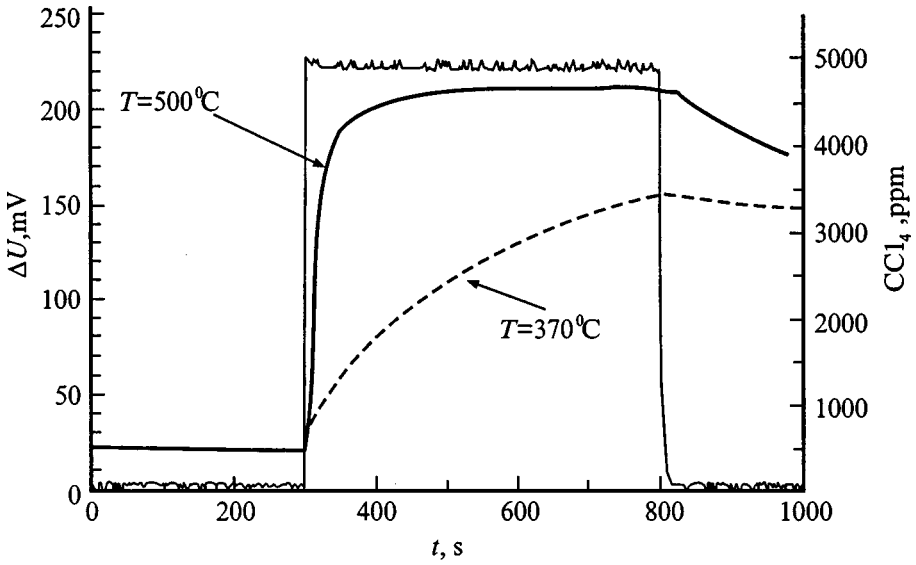

FIG. 8. Sensor signal after exposure of the structure to 5000 ppm tetrachloromethane in air. The concentration step of the investigated gas is shown in the figure.

trodes made from metals of the platinum group have a high sensitivity to hydrogen. However, an experiment with a Pt–SiO_2–6*H*-SiC structure has shown that no CH_4 response signal is generated at the given temperatures. This result convinces us that the dissociation of methane with the formation of hydrogen on the surface of Pt does not occur at the indicated temperatures. In the case of methane we are probably dealing once again with a change in the degree of oxidation of the subsurface layer of the LaF_3 due to a methane oxidation reaction.

If we compare typical chlorofluorocarbon relaxation curves (Fig. 4) with those for CCl_4 and CH_4, we find a significant difference in the rates of the leading and trailing edges. After the injection of pure air into the measurement chamber recovery of the initial bias voltage is so slow that the influence of these gases on the structure can hardly be deemed reversible (Figs. 8 and 9). Consequently, the successive action of CCl_4 and CH_4 concentration pulses on the structure eradicates the response. This effect on the structure can be regarded as a kind of initializing of the sensor, enhancing its selectivity.

Raising the working temperature of the sensor to 600 °C might diminish the response of the sensor to methane and is one of our future research goals. The increased rate of catalytic oxidation of methane on the surface of the platinum electrode could decrease the concentration of this gas reacting with the interphase boundary of the structure.

CONCLUSION

We have demonstrated the conceptual possibility of using a $Pt/LaF_3/SiO_2/SiC$ structure to diagnose the content of several chlorofluorocarbons in air. An increase in the working temperature enhances the sensitivity to Freons and improves the response time of the sensor. At a working temperature of approximately 500 °C it is possible to detect changes in the Freon concentrations at the 10-ppm level, and the time to attain a steady-state signal is approximately

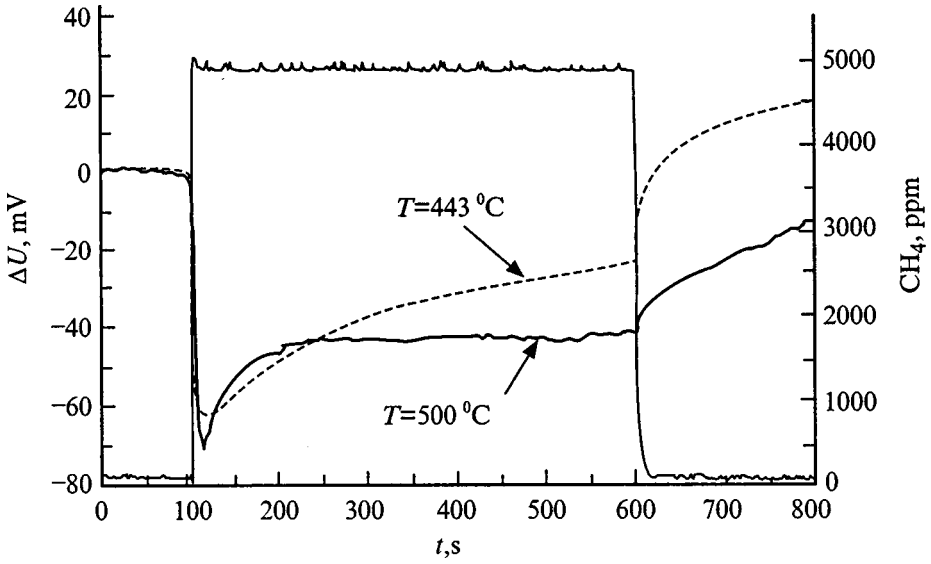

FIG. 9. Sensor signal after exposure of the structure to 5000 ppm methane in air. The concentration step of the investigated gas is also shown.

1 min. We have estimated the activation energy $\Delta\varepsilon$ from the temperature dependence $\partial\Delta U/\partial t|_{t=t_0}=f(T)$. This result brings us to the conclusion that the preliminary dissociation of Freon molecules could be the limiting stage of the gas-sensing process. However, it is evident from an estimate of the number of reaction electrons in the Nernst equation that the processes involved in the interaction of the structure with Freons depend on the temperature and have a different character than those occurring in the interaction with fluorine at temperatures up to 200°.

We have found that the structure is sensitive to tetrachloromethane and methane. The sensitivity of the structure to methane can be attributed to a change in the degree of oxidation of the solid electrolyte at the interphase boundary. Raising the working temperature and stabilizing the properties of the interphase boundary by, for example, presaturating it with fluorine have the effect of lowering the minimum detectable concentration and improving the selectivity.

This work has received support form the Volkswagen Foundation, Grant No. 1/70-753.

[1)]The semiconductor substrates were prepared at the A. F. Ioffe Physicotechnical Institute of the Russian Academy of Sciences in St. Petersburg, Russia.

[1] M. Shiratori, M. Katsyra, and T. Tsuchiya, in *Proceedings of the International Meeting on Chemical Sensors* (Fukuoka, Japan, 1983), pp. 119–124.
[2] T. Nomura, T. Amamoto, Y. Matsuura, and Y. Kajiyama, Sensors and Actuators B, Nos. 13–14, 486 (1993).
[3] W. Moritz, S. Krause, A. A. Vasiliev *et al.*, Sensors and Actuators B, Nos. 24–25, 194 (1995).
[4] W. Moritz, V. Filippov, and L. Bartolomäus *et al.*, in *Proceedings of the International Conference on Solid-State Sensors and Actuators* (Chicago, 1997), pp. 1073–1076.
[5] A. A. Terent'ev and V. I. Filippov, Zh. Anal. Khim. **52**, 635 (1997).
[6] A. A. Terent'ev, V. I. Filippov, and S. S. Yakimov, Poverkhnost', No. 2, 72 (1994).
[7] A. Arbab, A. Spetz, and I. Lundström, Sensors and Actuators B, Nos. 15–16, 19 (1993).

Translated by James S. Wood

Influence of the structure of the director field on the optical properties of a nematic liquid-crystal droplet

V. A. Loĭko and V. I. Molochko

B. I. Stepanov Institute of Physics, Academy of Sciences of Belarus, 220072 Minsk, Belarus
(Submitted May 19, 1998)
Zh. Tekh. Fiz. **69**, 86–90 (November 1999)

The changes in the optical properties of a spherical, bipolar, liquid-crystal droplet due to the change in structure of its director field under the influence of an applied external field are investigated on the basis of calculations using the nested-ellipsoids model in the discrete-dipole approximation. Graphs are plotted for the attenuation efficiency factor and the angular structure of the scattered radiation for the two typical cases of strong and weak dependence of the optical properties of the droplet on the change in structure of its director field.
 [S1063-7842(99)01511-1]

INTRODUCTION

The application of liquid crystals in the form of polymer-dispersed liquid crystal (PDLC) droplets in optical devices is attracting attention as a promising technique for the design of a number of PDLC-based electrooptical elements. The electrooptical parameters of such systems depend on their production technology.[1,2] They are determined by the size and shape of the droplets, the concentration and orientation of the droplets in the layer, the refractive indices of the polymer and the droplets, etc. The most significant factors affecting the properties of the PDLC layer, as a rule, are the orientation of the droplets in the sample and the greater or lesser extent to which the directors of the individual droplets in the PDLC layer are ordered. The degree of this order changes when an external electric field is applied to the sample, so that the optical properties of the sample itself change under the influence of the electric field. In the majority of cases this effect is far greater in magnitude than all others and is the only effect taken into account in calculations. On the other hand, in the presence of an applied electric field other changes influencing its optical properties are known to take place in a PDLC sample.[1–5] As a rule, to include these changes poses a fairly complex problem, and their contribution to the optical properties of the sample is not as pronounced as the contribution from the droplet orientation effect, and in most papers this disparity is exploited to describe the optical properties of droplets on the basis of simplified models.[6–12]

One effect is a transformation of the internal structure of the director field of the PDLC droplet under the influence of an applied external voltage. For a sample in zero field or in a weak electric field, when the directors of the droplets are disordered, the influence of this effect is obviously insignificant. In the transparency mode, on the other hand, when the droplet directors are more or less ordered under the influence of the field, its influence can be appreciable. For PDLC layers we encounter the problem of inadequate transparency, or turbidity, even in an applied external field.[1,7] In the transparency mode the influence of the indicated effect, which is inconsequential in other situations, can significantly increase or decrease the turbidity of the sample. In this case the external field-induced transformation of the structure of the director field of the PDLC droplet must be taken into account.

The present study is devoted to the investigation and assessment of the influence of transformation of the structure of the director field on the optical properties of a PDLC droplet in the example of the most commonly encountered case of a bipolar director structure. Our choice of analytical method is the discrete-dipole approximation (DDA), because the accuracy of conventional approximations (the Rayleigh–Gans and anomalous diffraction approximations) is not always sufficient for the description of scattering by a liquid-crystal droplet.

BASIC RELATIONS

We consider a liquid-crystal droplet imbedded in a polymer with a refractive index n_m in an incident light field $\mathbf{E}^{\text{inc}}=\mathbf{e}_0 E_0 \exp(i\mathbf{k}_0\cdot\mathbf{r})$, where $\mathbf{e}_0$ is the unit polarization vector. We denote by $\mathbf{k}$ the wave vector in the scattering direction, align the x axis of a Cartesian coordinate system with the direction of the droplet director, and denote the angle between $\mathbf{k}$ and the droplet director by α (Fig. 1). We write the scattered field in the usual form

$$\begin{pmatrix} E_{\parallel} \\ E_{\perp} \end{pmatrix} = \frac{\exp(ikr)}{-ikr} \begin{pmatrix} S_2 & S_3 \\ S_4 & S_1 \end{pmatrix} \begin{pmatrix} E_{\parallel}^{\text{inc}} \\ E_{\perp}^{\text{inc}} \end{pmatrix}, \tag{1}$$

where $E_{\perp}$ and $E_{\parallel}$ are the components of the vector $\mathbf{E}$ perpendicular and parallel to the scattering plane.

We use relative refractive indices below, setting $n_m=1$. We denote the wave number by $k\equiv 2\pi/\lambda$, where λ is the wavelength in the polymer matrix. We assume that the liquid crystal is uniaxial. In this case its dielectric tensor, reduced to diagonal form, has two identical values $\hat{\boldsymbol{\varepsilon}}=\text{diag}\,(n_o^2, n_o^2, n_e^2)$ (n_o and n_e are the refractive indices of the ordinary and extraordinary rays, respectively).

To calculate the optical properties of the liquid-crystal droplet, we use the discrete-dipole approximation, which is

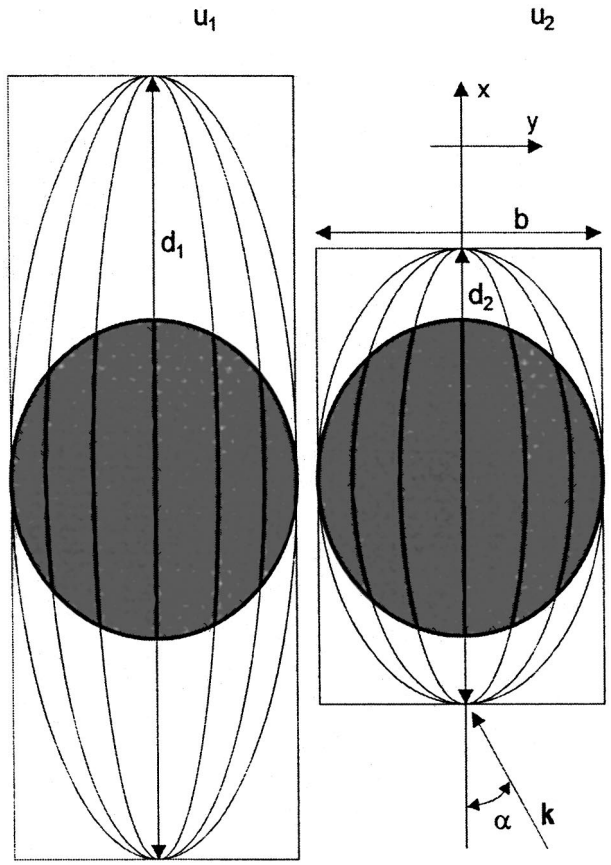

FIG. 1. Geometry of the nested-ellipsoids model for a liquid-crystal droplet with a bipolar structure.

based on representation of a scatterer by an array of small scatterers (dipoles), whose polarizability is determined from their size and the characteristics of the droplet medium. Each scatterer has its own characteristic position and polarizability, the choice of which poses a nontrivial problem.[13,14]

In accordance with Refs. 13 and 14 the problem of finding the polarization of each elementary dipole in the external field $\mathbf{E}^{\text{inc}}$ is reduced to solving the system of linear equations

$$\mathbf{P}_j=\hat{\boldsymbol{\alpha}}_j\left[\mathbf{E}^{\text{inc}}+\sum_{k\neq j}\hat{\mathbf{A}}_{jk}\mathbf{P}_k\right]. \tag{2}$$

Here $\hat{\boldsymbol{\alpha}}_j$ is the polarizability of the jth dipole (see the expressions for $\hat{\boldsymbol{\alpha}}_j$ in Refs. 13 and 14), and the term $\hat{\mathbf{A}}_{jk}\mathbf{P}_k$ represents the field induced at the jth dipole by the kth dipole with polarization $\mathbf{P}_k$.

As a rule, the solution of the system of equations (2) reduces to iterations with the physical significance that higher and higher multiplicities of scattering by the system of N elementary dipoles are included. For the polarizability of the jth dipole we have

$$\hat{\boldsymbol{\alpha}}_j=\hat{\mathbf{M}}_j^{-1}\hat{\boldsymbol{\alpha}}\hat{\mathbf{M}}_j, \tag{3}$$

where $\hat{\boldsymbol{\alpha}}$ is the diagonal polarizability tensor given in Refs. 13 and 14,

$$\hat{\mathbf{M}}_j=\hat{\mathbf{R}}_x(\alpha_j)\hat{\mathbf{R}}_y(\beta_j)\hat{\mathbf{R}}_z(\gamma_j), \tag{4}$$

α_j, β_j, and γ_j are the angles of rotation of the jth dipole about the x, y, and z axes, and $\hat{\mathbf{R}}_x(\alpha)$, $\hat{\mathbf{R}}_y(\beta)$, and $\hat{\mathbf{R}}_z(\gamma)$ are the matrices of rotation about these axes, respectively.

We use the nested-ellipsoids model to describe the field of the director in a particle. In this parametrization the liquid-crystal molecules in the droplet are oriented along a line lying in the same plane as the director of the droplet and tangential to the surfaces of a system of nested concentric ellipsoids with a common axis (Fig. 1).

In this case the polarizability tensor $\hat{\boldsymbol{\alpha}}_j$ for a spheroidal droplet is obtained by rotating the polarizability tensor $\hat{\boldsymbol{\alpha}}$ used in Refs. 13 and 14 through the angles

$$\alpha_j=\arctan(z_j,y_j), \tag{5a}$$

$$\beta_j=0, \tag{5b}$$

$$\gamma_j=\frac{\pi}{2}-\arctan(x_j\sqrt{y_j^2+z_j^2},(b(1+u))^2-x_j^2), \tag{5c}$$

$$u\equiv\frac{d-b}{b}, \tag{5d}$$

$$-b/2\leqslant x_j\leqslant b/2, \tag{5e}$$

$$-a/2\leqslant y_j\leqslant a/2, \tag{5f}$$

$$-a/2\leqslant z_j\leqslant a/2. \tag{5g}$$

Here a and b are the lengths of the axes of the ellipsoidal droplet, and d is the common axis of the system of nested ellipsoids modeling the director configuration in the droplet (Fig. 1); if $s=\sin(t)$ and $c=\cos(t)$, then $\arctan(s,c)=t$.

For simplicity we shall assume from now on that the droplet is spherical, i.e., $a=b$. The tensor $\hat{\alpha}$ is determined in accordance with Refs. 13 and 14 for the dielectric permittivity $\hat{\boldsymbol{\varepsilon}}=\text{diag}(\varepsilon_o,\varepsilon_e,\varepsilon_o)$.

The parameter u given by Eq. (5d) is equal to zero for a bipolar ellipsoidal droplet without an external field. As u tends to infinity, the bipolar droplet goes over to a droplet with uniformly oriented molecules. We shall regard the parameter u below as a model external voltage applied to a bipolar ellipsoidal droplet. The nested-ellipsoids model has been used previously[1,15–17] to describe the structure of a bipolar droplet.

For the scattering efficiency factor, which is equal to the attenuation efficiency factor in the case of nonabsorbing droplets, we have[13]

$$Q_{\text{sca}}=\left(\frac{4}{xE_0}\right)^2\sum_{j=1}^{N}\text{Im}\,(k^3\mathbf{P}_j\cdot\mathbf{E}_j^{\text{inc}}), \tag{6}$$

where $x\equiv\pi a/\lambda$ is the diffraction parameter of a spherical droplet.

In the ensuing calculations we characterize the angular structure of the scattered radiation by the function I:

$$I\equiv\frac{1}{2}(|S_1|^2+|S_2|^2). \tag{7}$$

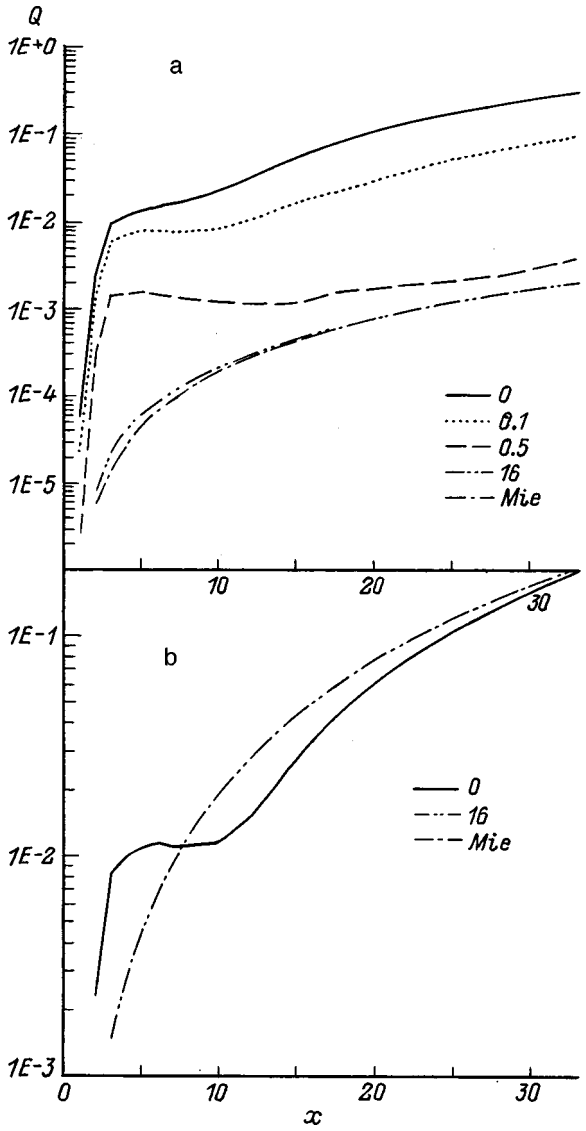

FIG. 2. Dependence of the attenuation efficiency factor Q of a spherical, bipolar, liquid-crystal droplet on the droplet diameter x for: $n_o=0.999$ (a) and $n_o=0.99$ (b), $n_e=1.15$, values of the parameter $u=0, 0.1, 0.5$, and 16 (a) and $u=0$ and 16 (b), and light incident in the same direction as the director of the droplet, $\alpha=0$. The dot-and-dash curves are calculated for a homogeneous Mie scatterer with refractive indices $n=0.999$ (a) and $n=0.99$ (b).

The dimensionless value of I is related to the differential scattering cross section $d\sigma/d\Omega$ for natural light by the equation

$$I=\frac{d\sigma}{d\Omega}k^2. \tag{8}$$

RESULTS OF THE CALCULATIONS

Figure 2 shows the results of calculations of the attenuation efficiency factor of a bipolar spherical droplet as a function of its diameter x and the model voltage u for $n_o=0.999$ (a) and 0.99 (b), $n_e=1.15$, and values of the voltage $u=0, 0.1, 0.5$, and 16 (a) and $u=0.16$ (b). The value $u=0$ corresponds to a bipolar scatterer in zero field, and $u=16$ essentially corresponds to a homeotropic crystal (codirectional axes of the liquid-crystal molecules).

The choice of $n_o=0.999$ for the ordinary refractive index leads to a strong dependence of the attenuation efficiency factor on the parameter u. For example, the factor differs by more than two orders of magnitude for $u=0$ and $u=16$. This means that even if the sample parameters are selected very precisely, the change in structure of the bipolar scatterer upon application of an electric field produces a significant variation or, on the whole, turbidity of the PDLC sample.

Figure 2a also shows the variation of the attenuation efficiency factor of a homogeneous droplet (Mie scatterer model) whose refractive index n is equal to the ordinary index n_o. The fact that this curve is very close to the $u=16$ curve indicates that the model of a homogeneous particle with an effective refractive index is applicable for describing the attenuation efficiency factor of a bipolar droplet in a high field in our situation. In Fig. 2b ($n_o=0.99$), on the other hand, a strong dependence of the attenuation efficiency factor on the model voltage u is not observed. The curves for $u=0$ and $u=16$ differ only slightly, and the Mie scattering curve and the $u=16$ curve coincide.

The difference in the patterns of behavior of the attenuation efficiency factor as a function of the parameter u in Fig. 2 can be attributed to different ratios between the contributions of two parts: scattering by the homogeneous effective scatterer and scattering by inhomogeneities of the refractive index (structure) of the scatterer. In Fig. 2b ($n_o=0.99$) the absolute value of the scattering efficiency factor is much higher than in Fig. 2a ($n_o=0.999$). On the other hand, the contribution of scattering by structural inhomogeneities of the droplet is approximately equal in the two cases (since it depends mainly on the extraordinary refractive index n_e and the structure of the director) and is of the order of 10^{-2} in Fig. 2. Whereas the value of this contribution is substantial against the background of a total attenuation efficiency factor of the order of $10^{-4}-10^{-3}$ in Fig. 2a, its influence is slight at the attenuation efficiency factor of the order of $10^{-2}-10^{-1}$ in Fig. 2b.

The graphs in Fig. 2 can be used to estimate the influence of the change in structure of a bipolar PDLC droplet on its optical properties. For the description of a PDLC sample the frequently encountered homeotropic droplet model (i.e., disregard for the influence of the droplet structure) describes its optical properties in most cases. On the other hand, for a sample in the transparency mode, when the directors of the particles are oriented along the field, the influence of the change in the structure of the director field in the presence of an external field can make a significant contribution. Moreover, it may be necessary to take this effect into account in order to solve the turbidity problem for PDLC layers[1] in the transparency mode.

Figures 3a and 3b show the variations of the angular structure of the scattered radiation for: $x=5$, the same refractive indices as in Figs. 2a and 2b, and values of the parameter $u=0, 0.1, 0.5, 2$, and 16 (a) and $u=0$ and 16 (b). Also shown in Fig. 3 are graphs of the angular scattering function

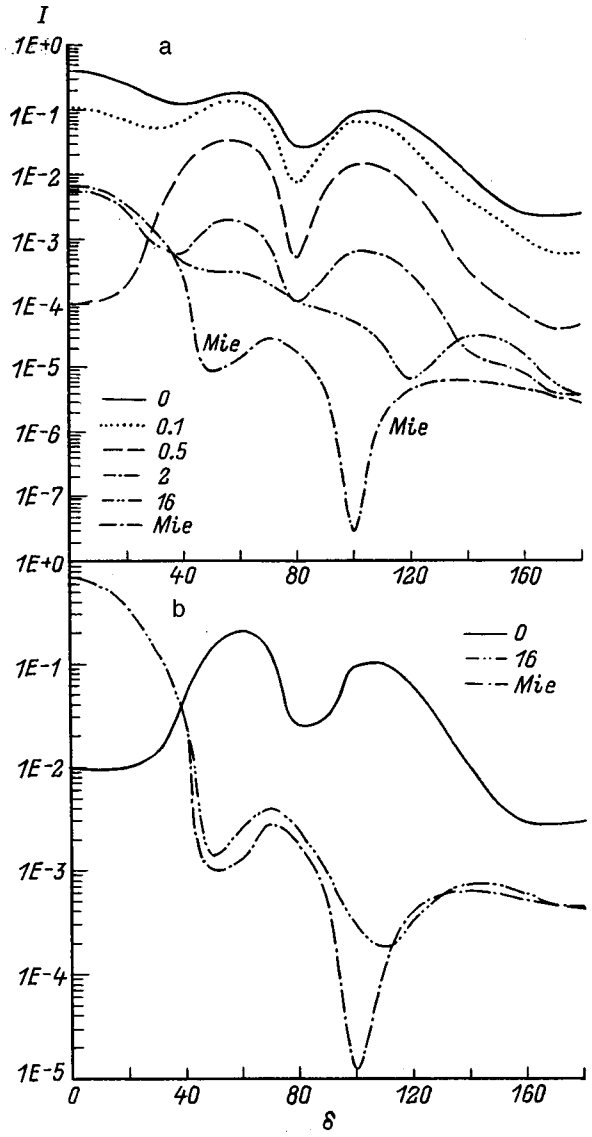

FIG. 3. Angular structure of scattered radiation $I=(|S_1|^2+|S_2|^2)/2$ from a spherical bipolar droplet of diameter $x=5$ for: $n_o=0.999$ (a) and $n_e=1.15$ (b), values of the parameter $u=0$, 0.1, 0.5, and 16 (a) and $u=0$ and 16 (b), and illumination by natural light incident in the same direction as the director of the droplet ($\alpha=0$). The dot-and-dash curves are calculated for a homogeneous Mie scatterer with refractive indices $n=0.999$ (a) and $n=0.99$ (b).

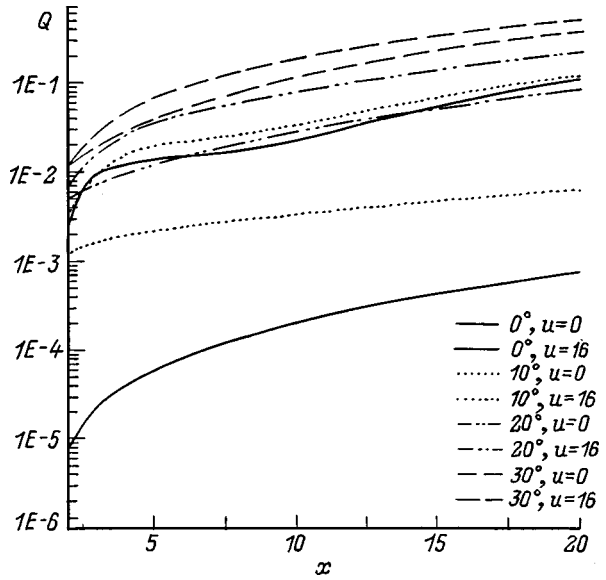

FIG. 4. Dependence of the attenuation efficiency factor Q of a spherical, bipolar, liquid-crystal droplet on the droplet diameter x for: $n_o=0.999$ and $n_e=1.15$, values of the model voltage $u=0$ and 16, and light at angles of incidence $\alpha=0°$, 10°, 20°, and 30°.

for a homogeneous particle having a refractive index equal to the ordinary index of the liquid crystal (Mie scatterer). A comparison with the graph for a homogeneous particle shows that for a PDLC droplet in a strong field the homeotropic droplet model well describes not only the total attenuation of the light beam, but also the angular structure of the scattering. On the other hand, for a bipolar structure and low ''voltages'' u the structure of the scattered radiation at large angles is dependent in either case on scattering by inhomogeneities of the refractive index of the droplet. This means that the configuration of the director must be taken into account in order to describe the angular structure of the scattered radiation, even in cases where the attenuation efficiency factor is satisfactorily described by the homeotropic model.

Figure 4 shows a graph of the scattering efficiency factor as a function of the droplet diameter for two model voltages u and four values of the angle α between the director of the droplet and the direction of incidence of the light. The parameters of the droplet in Fig. 4 correspond to the parameters of the droplets in Figs. 2a and 3a, $n_o=0.999$, and $n_e=1.15$. It is evident from the figure that the attenuation efficiency factor for a droplet in a strong field and without an external field differ by more than two orders of magnitude in our case for $\alpha=0$, whereas for $\alpha=30°$ the difference is approximately twofold. This stark contrast is attributable to the fact that when the angle α between the direction of the director of the droplet and the direction of light propagation increases for an ordinary refractive index chosen essentially to coincide with the refractive index of the medium, the droplet as a whole becomes more rigid, and the influence of the change in the structure of the director field on the optical properties of the droplet diminishes against the background of strong scattering.

CONCLUSION

We have attempted to assess how the optical properties are affected by the change in structure of the director field of a polymer-dispersed liquid crystal in the presence of an applied voltage. To model the applied external field, we have used the parameter u, which is equal to the ratio of the axes of a system of nested ellipsoids representing the direction of the director of the liquid crystal in the droplet. The relationship of this model voltage to the true voltage depends on the parameters of the PDLC layer, but certain conclusions can still be drawn on the basis of our simple analysis.

On the whole, the simple homeotropic model adequately describes the attenuation of light by a PDLC droplet. On the other hand, to describe a PDLC layer in the transparency mode and the angular structure of the scattered radiation, it

may be essential to take into account the configuration of the director in the droplet and the change in configuration under the influence of an external voltage.

This investigation has received partial support from the U.S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF Grant N BE1-113).

[1] P.S. Drzaic, in *Liquid Crystal Dispersions* (World Scientific, Singapore, 1995), 430 pp.
[2] G. M. Zharkova and A. S. Sonin, in *Liquid-Crystal Composites* [in Russian], Nauka, Moscow (1994), 213 pp.
[3] A. V. Koval'chuk, M. V. Kurik, O. D. Lavrentovich, and V. V. Sergan, Zh. Éksp. Teor. Fiz. **94**(5), 350 (1988) [Sov. Phys. JETP **67**, 1065 (1988)].
[4] J. H. Erdman, S. Zumer, and J. W. Doane, Phys. Rev. Lett. **64**, 1907 (1990).
[5] R. Ondris-Crawford *et al.*, J. Appl. Phys. **69**, 6380–6386 (1991).
[6] P. S. Drzaic, in *SPIE*, Vol. 1911, *Liquid Crystals Chemistry. Physics and Applications*, 1993, pp. 153–159.
[7] P. S. Drzaic, in *SPIE*, Vol. 2175, *Liquid Crystals Chemistry. Physics and Applications*, 1994.
[8] B.-G. Wu, J. L. West, and J. W. Doane, J. Appl. Phys. **62**, 3925 (1987).
[9] O. A. Aphonin, Yu. V. Panina, A. B. Pravdin, and D. A. Yakovlev, Liq. Cryst. **15**, 395 (1993).
[10] P. S. Drzaic, Mol. Cryst. Liq. Cryst. **261**, 383 (1995).
[11] J. R. Kelly and P. Palffy-Muhoray, Mol. Cryst. Liq. Cryst. **243**, 11 (1994).
[12] F. Bloisi, P. Terrecuso, L. Vicari, and F. Simoni, Mol. Cryst. Liq. Cryst. **266**, 229 (1995).
[13] B. T. Draine and P. J. Flatau, J. Opt. Soc. Am. **11**, 1491 (1994).
[14] B. T. Draine, Astrophys. J. **333**, 848 (1988).
[15] R. D. Williams, J. Phys. A **19**, 3211 (1986).
[16] F. Xu, H. S. Kitzerow, and P. P. Crooker, Phys. Rev. E **49**, 3061 (1994).
[17] J. Ding and Y. Yang, Jpn. J. Appl. Phys. **31**, 2837 (1992).

Translated by James S. Wood

Controlled diffractive optical elements containing a vanadium dioxide film

O. B. Danilov and A. I. Sidorov

Scientific-Research Institute of Laser Physics, 199034 St. Petersburg, Russia
(Submitted August 24, 1998)
Zh. Tekh. Fiz. **69**, 91–96 (November 1999)

The principles of constructing optical devices on the basis of diffractive structures with a vanadium dioxide film for the control of radiation in the mid-IR range are analyzed. Methods are described for the practical implementation of such devices at $\lambda = 10.6\ \mu$m, and their response characteristics are calculated. It is shown that a contrast of $1:10^7$ can be attained in diffractive optical elements, and the actuation time of the elements when switched on by an intense laser beam can be shortened to tens of nanoseconds. © *1999 American Institute of Physics.* [S1063-7842(99)01611-6]

INTRODUCTION

The recent fascination with diffractive optical elements (DOEs) derives mainly from the encouraging outlook for their application in optical signal processing and imaging systems, including computer optics. Apart from the traditional use of DOEs as spectral selectors, in recent times we have seen the development of a sizable assortment of different types of DOEs capable of implementing many other functions: beam multiplexing[1] and shaping,[2] the allocation of optical signals to processing channels,[3] wave-front shaping,[4] etc. The optical characteristics of such devices can be fixed, or they can be controlled by the introduction of electrooptical materials.[5,6] The above-indicated domain of application of DOEs is the reason that the majority of DOE research and development is conducted in the visible and near-IR ranges.

A whole series of problems is also encountered in connection with the processing and switching of optical signals in mid-IR laser optical systems. Examples can be found in the problems of separating weak and strong signals in lidar systems, the distribution of signals among photodetection devices, and the temporal and spatial selection of signals. One of the most important problems is the protection of photodetection devices against damage and ''blinding'' by intense laser beams.

Multilayer interference (MLI) systems utilizing a VO_2 film as a control element are widely used for the modulation of radiation in the mid-IR range. The reversible semiconductor–metal transition[7] that takes place in a single-crystal or polycrystalline VO_2 film as the temperature varies is accompanied by significant variations of its optical constants. This behavior can be exploited to modify the optical characteristics of an interference system utilizing a VO_2 film.

Multilayer interference systems containing a VO_2 film can be fabricated as transmission[8] or reflection[9] devices. In this paper we discuss only interference systems of the reflection type — VO_2 mirrors. This choice is dictated by two considerations. First, VO_2 mirrors can invest DOEs with broader functional capabilities. For example, in MLI systems of the transmission type the transmission can only decrease as the temperature is increased ($dA/dT<0$, where A is the transmissivity). In MLI systems of the reflection type the derivative dR/dT can be positive, negative, or an alternating function, depending on the type of interference system,[10] and phase modulation of the reflected radiation can be achieved with a small variation of the reflectivity R (Ref. 11). In the second place, systems of the reflection type containing VO_2 films have a much higher resistance to beam damage than do systems of the transmission type and can therefore be used to control high-intensity radiation.

In this paper we discuss the principles of construction of controlled DOEs with a VO_2 film in the spectral range $2.5-12\ \mu$m. The principal application of such DOEs is in the control of high-intensity chemical, CO, and CO_2 laser beams. Our analysis is therefore concerned primarily with methods for achieving maximum contrast while preserving the high beam stability of DOEs. The main results are given for $\lambda = 10.6\ \mu$m (the emission wavelength of a CO_2 laser).

PRINCIPLES OF CONSTRUCTION OF CONTROLLED DIFFRACTIVE OPTICAL ELEMENTS WITH A VO_2 FILM

Controlled DOEs that utilize MLI systems of the reflection type with VO_2 films constitute interference systems in which a spatial structure is formed (Fig. 1a) from alternating regions with constant and temperature-dependent optical parameters. The regions of the first type are characterized by the reflectivity R_1 and by the phase φ_1 of the reflected beam, while the corresponding parameters for the regions of the second type are $R_2(T)$ and $\varphi_2(T)$. Figures 1b–1d show the simplest DOEs of this type: a constant-period diffraction grating (b), a one-dimensional Fresnel zone plate (c), which can be used to focus radiation into a line, and a Fresnel concentric zone plate (d), which can be used to focus radiation into a point (or spot). The figures show only the first diffraction order. The function of the mirrors M1 in Figs. 1c and 1d is to spatially separate the incident and diffracted beams; the central Fresnel zones are not involved in the diffraction process. For the investigated DOE structures and a plane wave at normal incidence the radiation intensity in a diffraction order can be written in the form

$$I = I_0 K(R_1 + R_2 - 2\sqrt{R_1 R_2}\cos(\varphi_1 - \varphi_2)). \quad (1)$$

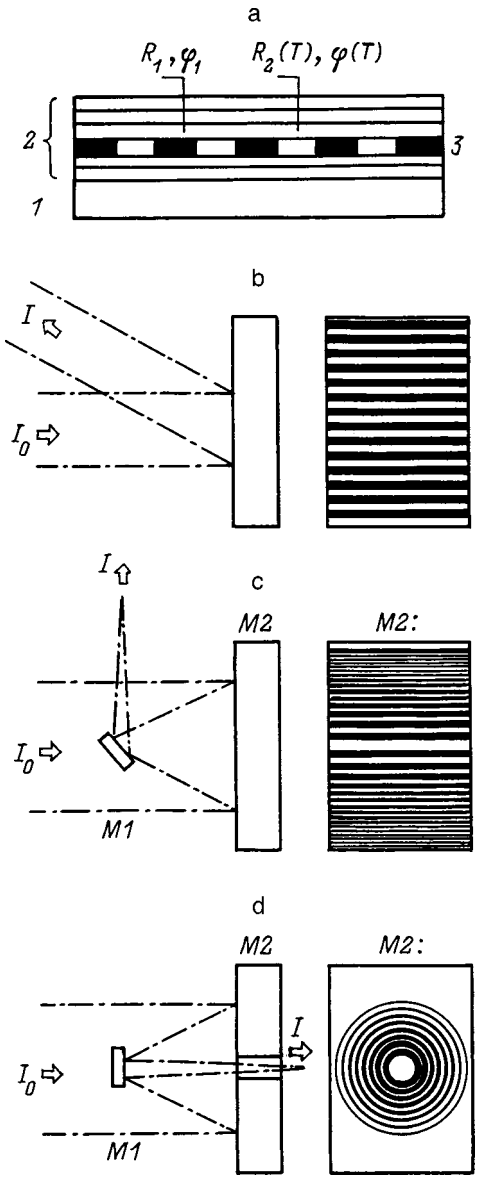

FIG. 1. a: Structure of a DOE with a VO_2 film: *1* — substrate; *2* — multilayer interference (MLI) system; *3* — periodic structure with a VO_2 film; b–d: Various DOE configurations.

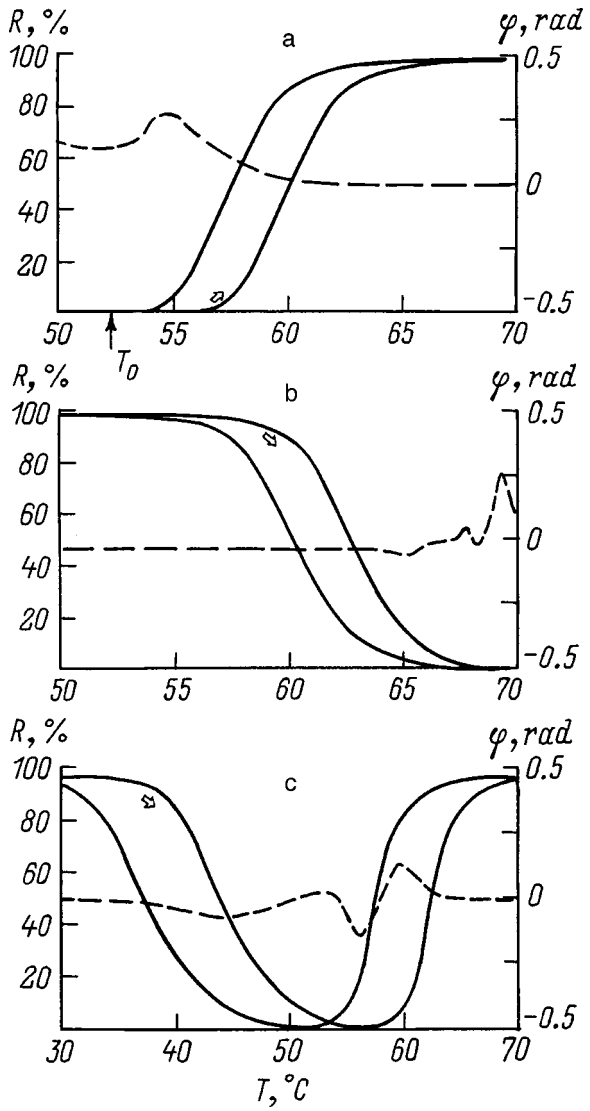

FIG. 2. Temperature dependences of the reflectivity $R(T)$ (solid curves) and the phase $\varphi(T)$ (dashed curves) of the reflected radiation for VO_2 mirrors.

Here I_0 is the intensity of the incident beam, and K is a parameter that depends on the geometry of the diffraction structure and the diffraction order. It follows from this equation that the contrast in a controlled DOE can be maximized if one of the states of the DOE corresponds to a mirror with a uniform reflectivity and a reflected beam with a uniform phase: $R_1 = R_2$ and $\varphi_1 = \varphi_2$. In this case all the incident radiation is specularly reflected, and the radiation intensity in the diffraction orders is equal to zero. The situation of greatest practical interest is when the reflectivity R_1 is equal to the maximum reflectivity, $R_2(T) = R_{2\max}$. The maximum diffraction efficiency in the diffraction orders is achieved when either of two conditions is satisfied: $R_1 \gg R_2$ or $\varphi_1 - \varphi_2 = \pi$ and $R_1 = R_2$.

Figure 2 shows calculated temperature curves of the reflectivity and the phase of the reflected beam for the three main types of VO_2 mirrors suitable for the fabrication of controlled DOEs. Only the ascending-temperature branches of the temperature hysteresis loop are shown for the $\varphi(T)$ curves. The curves shown here correspond to the following interference system configurations: (a) Ge(0.66)–BaF_2(1.96)–VO_2(0.3)–ZnS(0.815)–Al(0.1)–substrate, $dR/dT > 0$, $R_{\min} = 1\%$, $R_{\max} = 97.5\%$; (b) Ge(0.66)–ZnS(2.435) – VO_2(0.25) – ZnS(1.26) – Al(0.1) – substrate, $dR/dT < 0$, $R_{\max} = 98\%$, $R_{\min} = 0.1\%$; (c) Ge(0.66)–ZnS(0.87) – VO_2(0.25) – ZnS(1.26) – VO_2^*(0.25) – ZnS(1.26)–Al(0.1)–substrate, alternating dR/dT, $R_{\max}(30\,^\circ\mathrm{C}) = 98\%$, $R_{\min}(56\,^\circ\mathrm{C}) = 0.4\%$, $R_{\max}(70\,^\circ\mathrm{C}) = 94\%$. The numbers in parentheses give the thicknesses of the films in micrometers. The optical characteristics of the listed interference systems are calculated from the Fresnel equations by a recursive method. The temperature dependences of the refractive indices and the absorption of the stoichiometric VO_2 film[11] are used in the calculations. The configuration (c) contains two VO_2 films, one of which (VO_2^*) has the phase transition shifted to lower temperatures ($T = 30 - 50\,^\circ\mathrm{C}$) as a result of doping.[7,12]

It has been shown above that the maximum contrast is attained under the conditions $R_1 = R_{2\max}$ and $\varphi_1 = \varphi_2$. It is important to inquire, therefore, how closely the quantities $\Delta R = R_1 - R_{2\max}$ and $\Delta\varphi = \varphi_1 - \varphi_2$ approach zero under the condition $R_1 = \text{const}$. One possible way to form regions with a constant reflectivity is by inserting a thin metal film into the MLI system. An analysis shows, however, that only one parameter — either ΔR or $\Delta\varphi$ — can be minimized in this case. Both ΔR and $\Delta\varphi$ can be minimized simultaneously by introducing at least two additional layers into the MLI system. As an example, we consider the configuration (1), in which regions having a constant reflectivity are formed by two thin metal films situated on opposite sides of the VO_2 film: $Ge(0.642)-BaF_2(1.96)-Ti(0.0044)-VO_2(0.3)-Au(0.015)-ZnS(0.815)-Al(0.1)$. In this configuration the Au film provides imparts a high reflectivity to the system before the onset of the phase transition in the VO_2 film. The main purpose of the TiF Film is to equalize the phase of the reflected beam between the regions of constant and variable optical characteristics after the completion of the phase transition. Additional phase matching is achieved by slightly decreasing the thickness of the Ge film. For a DOE incorporating the above-described structure we have $\Delta R(70\,^\circ\text{C}) = 4\times10^{-5}\,\%$, and $\Delta\varphi(70\,^\circ\text{C}) = 5.5\times10^{-4}$ rad. During the phase transition in the VO_2 film the reflectivity R_1 varies by approximately 1%, which scarcely affects the optical characteristics of the DOE. A similar result can be obtained for the VO_2 mirror configurations (b) and (c) as well.

It is evident form Fig. 2 that the phase of the reflected beam varies considerably near the minimum of the reflectivity. Inasmuch as $R_{\min} \ll R_{\max}$ for the investigated VO_2 mirror configurations, DOEs utilizing them closely approximate pure amplitude elements, and their diffraction efficiency is no more than a few percent. The diffraction efficiency can be raised to 10% by combining a DOE with a Fabry–Perot etalon and choosing a ''resonance'' angle of incidence of radiation on the DOE.[13,14]

Another possible approach to improving the diffraction efficiency of DOEs is to use VO_2 mirrors that provide phase modulation of the radiation with a small variation of the reflectivity.[11] Figure 3 shows the phase and amplitude responses of such a VO_2 mirror with the following MLI system configuration: $BaF_2(4.522)-VO_2(0.25)-Al(0.1)-$substrate (the ascending-temperature branches of the temperature hysteresis loop are shown in the figure). This MLI system consists of a Gires–Tournois interferometer,[15] in which the VO_2 film functions as a control element. Regions with R_1 and φ_1 are formed in such a system by varying the thickness of the BaF_2 film. For example, a reduction of the thickness of the BaF_2 film to 4.51 μm produces the $\varphi(T)$ curve *2* in Fig. 3. The $R(T)$ now coincides with the $R(T)$ curve for the above-described structure within 0.1% error limits. A phase DOE constructed on the principle discussed here has the following characteristics: $\Delta\varphi(55\,^\circ\text{C}) \approx \pi$, $\Delta\varphi(70\,^\circ\text{C}) = 0.015$ rad.

The DOE can be controlled by means of an electron beam,[16] a thin-film heater[17] inserted between the MLI system and the substrate, or an intense laser beam. The second and third techniques are the most intriguing from the practical standpoint in that, unlike the electron-beam technique, they afford the capability of heating a large area of the DOE surface. Experimental and analytical results[18] show that when the VO_2 mirror is controlled by a thin-film heater, the minimum actuation time of the mirror is limited by the dielectric breakdown strength of the MLI system and is equal to $1-5\,\mu$s. The minimum actuation time of laser-controlled VO_2 mirrors[18] is limited by the resistance of the mirrors to high-intensity beam damage. For VO_2 mirrors with $dR/dT > 0$ the damage threshold is $2-5$ MW/cm^2 for a laser pulse of duration shorter than 1 μs. For a beam intensity ~ 1 MW/cm^2 the actuation time of the VO_2 mirrors is 10–100 ns, depending on the type of VO_2 mirror, the thickness of the MLI system, and the substrate material.

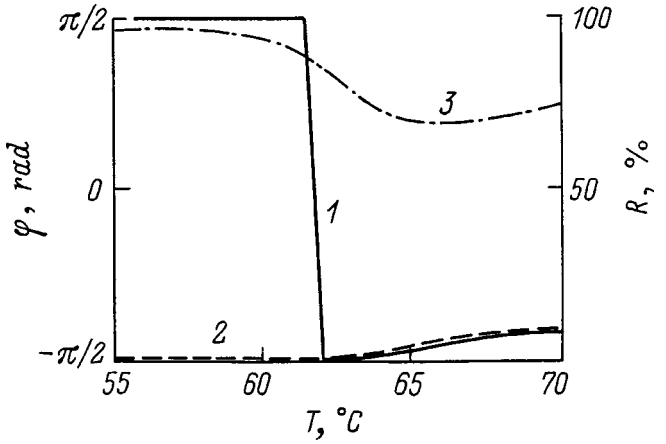

FIG. 3. The same as Fig. 2 for a phase DOE: *1* — $\varphi(T)$ for regions of the DOE with φ_2; *2* — $\varphi(T)$ for regions of the DOE with φ_1; *3* — $R(T)$ for regions of the DOE with R_1 and R_2.

The decay time of the VO_2 mirrors depends on the rate of heat transfer from the MLI system into the substrate by heat conduction and can have values of $10-50\,\mu$s for metal substrates. Since the actuation and decay of DOEs utilizing VO_2 mirrors are governed by a thermal mechanism, their application for the control of high-intensity radiation is limited to pulsed and periodic-pulse operation.

STATIC RESPONSE CURVES OF DIFFRACTIVE OPTICAL ELEMENTS WITH A VO_2 FILM

The modulation characteristics of amplitude DOEs were analyzed for $\Delta R = 10^{-3}\%$ and $\Delta\varphi = 5\times10^{-4}$ rad. Figure 4 shows the temperature dependence of the radiation intensity I in a diffraction order, normalized to the maximum radiation intensity $I_{\max}$ in the given diffraction order. It is evident that the radiation intensity in the diffraction orders decreases as the temperature increases in a DOE utilizing a VO_2 mirror with $dR/dT > 0$ (Fig. 4a, curve *1*; the temperature dependence of the reflectivity of this mirror is shown in Fig. 2a), and it becomes equal to $4\cdot10^{-7} I_{\max}$ at the maximum values of ΔR and $\Delta\varphi$.

In a DOE utilizing a VO_2 mirror with $dR/dT < 0$ (Fig. 2b) the radiation intensity in the diffraction orders decreases as R_2 decreases with increasing temperature. When an interference system with an alternating derivative dR/dT is used in the DOE (Fig. 2c), the radiation intensity in the diffraction

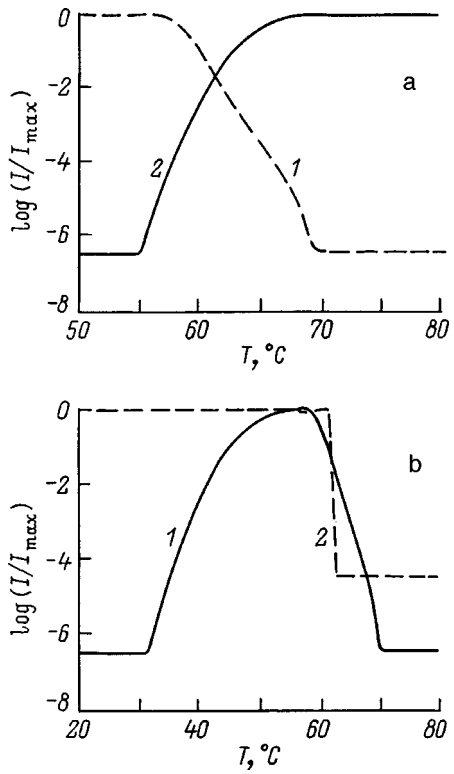

FIG. 4. Static responses of DOEs with VO_2 mirrors. a: *1* — MLI system with $dR/dT>0$; *2* — with $dR/dT<0$. b: *1* — alternating dR/dT; *2* — $d\varphi/dT<0$.

orders reaches a maximum in the temperature interval 50 $-55\,^{\circ}$C, after which it decreases to the initial level (Fig. 4b, curve *1*). For a DOE with a VO_2 phase mirror (Fig. 3) an increase in the temperature to 70° causes the radiation intensity in the diffraction orders to drop to $\sim 5\times 10^{-5} I_{\max}$ (Fig. 4b, curve *2*).

Thus the switching on and off of radiation in the diffraction orders can be achieved in controlled DOEs with VO_2 mirrors, and these two functions can be combined in a single DOE. For amplitude DOEs the derivative dI/dT has the opposite sign from the derivative dR/dT of the VO_2 mirror. For the investigated phase DOE the derivative dI/dT has the same sign as the derivatives $d\varphi/dT$ and dR/dT of the VO_2 phase mirror.

DYNAMIC RESPONSE CURVES OF DIFFRACTIVE OPTICAL ELEMENTS WITH A VO_2 FILM

The actuation-time dynamics in the switching of a DOE by a pulsed heat source depends on its power, type, and thermophysical parameters of the DOE and on the initial temperature of the VO_2 film (T_0 in Fig. 2a). The operation of greatest practical interest is the switching of DOEs driven by strong pulsed radiation, because the actuation time can be minimized in this case.

The VO_2 mirrors discussed in this paper have zero transmissivity (owing to the presence of a high-reflectivity metal film in the MLI system configuration). Consequently, the specific power of heat release in the MLI system with radiation heating can be written in the form

$$P = I_0(1-R). \tag{2}$$

The DOE with a VO_2 mirror contains regions whose reflectivity depends on the temperature. This dependence makes the specific power of heat release in these regions dependent on the temperature as well when the DOE is heated by radiation in the temperature interval of the phase transition in the VO_2 film: $P(T)=f[R(T)]$.

The actuation-time dynamics of a DOE driven by a radiation pulse can be analyzed on the basis of the approximation of the thermal model of a thin slab (MLI system) in ideal thermal contact with a semi-infinite body (substrate). For a heating time $t<500$ ns and a diffraction structure with a period typical of the mid-IR range, heat flow along the surface of the DOE by heat conduction can be disregarded. The temperature variation of the MLI system at a point $z\leqslant h$ (h is the thickness of the MLI system) in heating by a radiation pulse during a time Δt can be written in the form[19]

$$\Delta T(z)=\frac{1}{\sqrt{\pi}\eta_1}\left\{\sum_{n=0}^{\infty} M^{n+1}\int_0^{\Delta t}\frac{P(T)}{\sqrt{\Delta t-\tau}}\times\exp\left[-\frac{(2nh+h-z)^2}{4a_1(\Delta t-\tau)}\right]d\tau+\sum_{n=0}^{\infty} M^n\times\int_0^{\Delta t}\frac{P(T)}{\sqrt{\Delta t-\tau}}\exp\left[-\frac{(2nh+h+z)^2}{4a_1(\Delta t-\tau)}\right]d\tau\right\},$$

$$M=\frac{1-K}{1+K},\quad K=\frac{\eta_1}{\eta_2}. \tag{3}$$

Here the subscripts 1 and 2 refer to the slab and the semi-infinite body, respectively, $\eta=\sqrt{cdk}$, c is the specific heat, d is the specific gravity, k is the thermal conductivity, and a is the thermal diffusivity. The actuation-time dynamics of the DOE is modeled for the following conditions: $dR/dT>0$, Ge substrate, square radiation pulse with a uniform intensity distribution and $I_0=1$ MW/cm^2, and $T_0=55\,^{\circ}$C. Thickness-averaged thermophysical parameters of the MLI system are used in the calculations.

The time dependence of the radiation intensity in a diffraction order of a DOE with radiation pulse control is shown in Fig. 5 (curve *1*). It is seen that the intensity drops to the level $10^{-5} I_{\max}$ in 25 ns, after which the rate of change of the intensity decreases. This behavior is attributable to the fact that the DOE at this time is already in a state corresponding to a mirror with a high reflectivity, so that the fraction of radiation energy absorbed by the mirror and spent in heating it, decreases. The transient process is completely finished at $t=50$ ns. A decrease in the temperature T_0 has the effect of delaying the actuation of the DOE. For example, at $T_0=45\,^{\circ}$C the transient actuation of the DOE begins 15 ns after initiation of the radiation pulse.

For a DOE utilizing a Fresnel zone plate (Fig. 1d) the contrast can be further enhanced by employing a VO_2 mirror with $dR/dT<0$ as the mirror M1. This effect is illustrated by curve *2* in Fig. 5 for the case when M1 is situated in the plane corresponding to $\sqrt{2}$-fold constriction of the beam after reflection from the zone plate. The mirror M1 is actuated

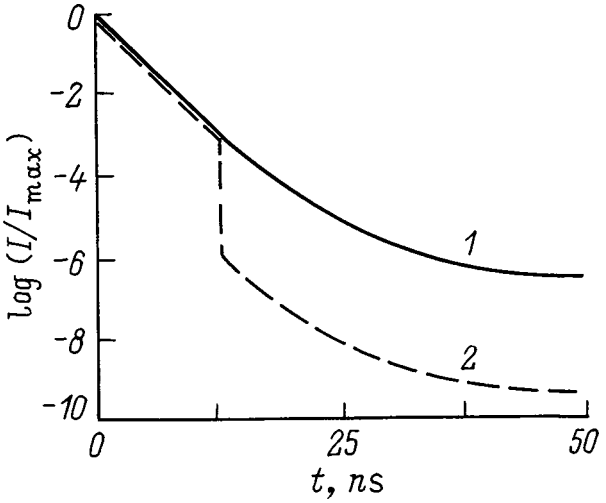

FIG. 5. Dynamic responses of a DOE utilizing: *1* — a VO_2 mirror, $dR/dT<0$, $I_0=1$ MW/cm^2; *2* — a Fresnel zone plate, $dR/dT<0$ (serving as mirror M1).

at $t=11$ ns in an avalanche manner as a result of the increase in the actuation rate as its reflectivity decreases. The minimum radiation intensity attained at the output of the given optical system is $\sim 5\times 10^{-10} I_{\max}$. It is obvious that even higher contrast can be achieved if the mirror M1 is a DOE utilizing a VO_2 mirror with $dR/dT>0$.

It is interesting to consider the actuation-time dynamics of a DOE with a one-dimensional periodic structure (Figs. 1b and 1c) when an initial temperature gradient along the ''rulings'' is specified in the VO_2 film, i.e., when $dT_0/dx=0$ (the x axis is parallel to the ''rulings'' of the structure). This technique causes regions of the DOE with a lower temperature T_0 to be actuated later than regions having a higher temperature. Figure 6a shows spatial distributions of the radiation intensity in a diffraction order of a DOE utilizing a VO_2 mirror with $dR/dT<0$ for the following conditions; $dT_0/dx=-5$ K/cm, $T_0(x=0)=55\,^\circ$C, $I_0=0.1$ MW/cm^2, copper substrate. At the initial time a uniform intensity distribution with $I_0=5\times 10^{-7} I_{\max}$ exists in the diffraction order. With increasing time t a region with the intensity $I_{\max}$ emerges in the vicinity of $x=0$ and subsequently spreads in the direction of $x>0$. Under the given conditions, therefore, a ''gate-opening'' function is implemented. The rate of travel of the actuation wave under these conditions is ~ 30 mm/μs.

Figure 6b shows the variation of the intensity in a diffraction order of a DOE utilizing a VO_2 mirror with an alternating derivative dR/dT. Modeling is performed for the same conditions as in the preceding case except that now $I_0=1.5$ MW/cm^2. It is evident from the figure that a narrow region of high intensity is formed in the diffraction order and moves toward increasing x with the passage of time. The constriction of the spatial distribution of the intensity about the maximum is associated with the high rate of change of the reflectivity at values close to $R_{\min}$. This case corresponds to the implementation of a ''moving slit'' function. The slit travels at a rate of ~ 160 mm/μs.

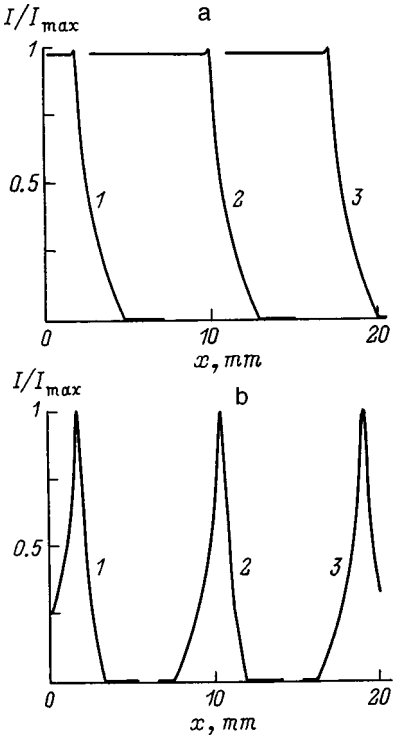

FIG. 6. Spatial distribution of radiation in a diffraction order for one-dimensional DOEs. a: MLI system with $dR/dT<0$ at: $t=260$ ns (*1*); 500 ns (*2*); 730 ns (*3*); b: MLI system with alternating dR/dT at: $t=50$ ns (*1*); 100 ns (*2*); 150 ns (*3*).

CONCLUSION

The results show that diffractive optical elements containing controlled interference systems of the reflection type based on VO_2 films can be used to make optical switching devices for the mid-IR range which exhibit the attributes of high contrast, fast response, and extensive functional capabilities. The principal application of such devices is for the control of intense laser beams in pulsed and periodic-pulse laser optical systems, including the separation of weak and strong signals, time selection, spatial scanning of radiation, and the protection of photodetection devices against high-intensity radiation damage.

[1] D. Mendlovic, Z. Zalevsky, G. Shabtay *et al.*, Appl. Opt. **35**, 6875 (1996).
[2] T. Dresel, M. Beyerlein, and J. Schwinder, Appl. Opt. **35**, 6865 (1996).
[3] K.-H. Brenner and F. Sauer, Appl. Opt. **27**, 4251 (1988).
[4] R. Piestun and J. Shamir, Opt. Lett. **19**, 771 (1994).
[5] T. Tatebayashy, T. Yamamoto, and H. Sato, Appl. Opt. **31**, 2770 (1992).
[6] Q. W. Song, X.-M. Wang, and R. Bussjager, Appl. Opt. **35**, 7031 (1996).
[7] A. A. Bugaev, B. P. Zakharchenya, and F. A. Chudnovskiĭ, *Metal–Semiconductor Junction and Its Applications* [in Russian], Nauka, Leningrad (1979), 183 pp.
[8] H. Jerominek, F. Picard, and D. Vincent, Opt. Eng. (Bellingham) **32**, 2092 (1993).
[9] A. B. Welch, B. Burzlaff, and W. Cunningham, Proc. Soc. Photo-Opt. Instrum. Eng. (SPIE) **300**, 153 (1981).
[10] O. P. Konovalova, A. I. Sidorov, and I. I. Shaganov, Opt. Zh. **65**, 20 (1998).
[11] O. P. Konovalova, A. I. Sidorov, and I. I. Shaganov, Opt. Zh. **65**, 24 (1998).
[12] M. Tazawa, P. Jin, and S. Tanemura, Appl. Opt. **37**, 1858 (1998).
[13] S. A. Collins and H. J. Caulfield, J. Opt. Soc. Am. A **6**, 1568 (1989).
[14] J. Krawczak, R. Dean, E. J. Torok *et al.*, Opt. Lett. **15**(22), 1264 (1990).

[15] F. Gires and P. Tournois, C. R. Acad. Sci. URSS **258**, 6112 (1964).
[16] J. S. Chivian, M. W. Scott, W. E. Case *et al.*, IEEE J. Quant. Electron. **QE-21**, 383 (1985).
[17] O. B. Danilov, O. P. Konovalova, A. I. Sidorov *et al.*, Pribory Tekh. Éksp., No. 4, 121 (1995).
[18] O. B. Danilov, V. V. Danilov, A. I. Sidorov *et al.*, in *Photonics West 1998, International Conference, San Jose, Calif., January, 1998* (SPIE, 1998).
[19] B. A. Grigor'ev, *Pulsed Radiation Heating* [in Russian], Vol. 2 (Nauka, Moscow, 1974), 727 pp.

Translated by James S. Wood

Angular dependence of the spin-wave resonance spectra in multilayer films

A. M. Zyuzin, A. G. Bazhanov, and V. V. Radaĭkin

N. P. Ogarev Mordovian State University, 430000 Saransk, Russia
(Submitted March 2, 1998)
Zh. Tekh. Fiz. **69**, 97–101 (November 1999)

The angular dependence of the spin-wave resonance spectra is investigated in multilayer magnetic films in which the spin pinning mechanism changes from dynamical to dissipative. A fundamental difference is observed in the character of the angular dependences of the spectra for three-layer films as opposed to two-layer films. In particular, the number of peaks in the spectrum of three-layer films is observed to decrease and then increase twice as the angle between the magnetic field and the normal to the film is varied by $\pi/2$. © *1999 American Institute of Physics.* [S1063-7842(99)01711-0]

Despite the vast number of papers devoted to spin-wave resonance (SWR) in thin films (see, e.g., Refs. 1–4), the angular dependence of the SWR spectra has not received nearly the attention that it should. In virtually every paper the analysis and experimental investigation of SWR spectra are carried out for perpendicular or parallel orientation of the static magnetic field $\mathbf{H}$ relative to the plane of the film. On the other hand, the analysis of the angular dependence can provide a better understanding of the properties and distinctive characteristics of the excitation of spin waves in magnetic films. Moreover, the variation of the angle between $\mathbf{H}$ and the film can, for certain values of the layer parameters, serve as a technique for smoothly varying the degree of spin pinning. The meager attention given to the angular dependences of the SWR spectra is largely attributable to the complexity of their interpretation. Of all the papers known to us, we can cite two[5,6] in which it is demonstrated experimentally how the SWR spectra evolve as the angle Θ_H between the field $\mathbf{H}$ and the normal to the film is varied. In these papers, however, the results are given only for two-layer films with close values of the damping parameter in the layers.

The objective of the present study has been to investigate the angular dependences of the SWR spectra in two-layer and three-layer films having a broad range of magnetic parameters.

The investigations were carried out on single-crystal iron garnet films prepared by liquid-phase epitaxy on (111)-oriented gadolinium–gallium garnet substrates. Two-layer and three-layer films were obtained by successive immersion in two or three different solutions in the melt. The thicknesses of the layers h were measured by the interference method, and the saturation magnetization $4\pi M$ was measured by a procedure described in Ref. 7. The effective uniaxial anisotropy field $\mathbf{H}_{ku}^{\text{eff}}=\mathbf{H}_{ku}-4\pi\mathbf{M}$ (where $\mathbf{H}_{ku}$ is the growth component of the uniaxial anisotropy field) and the gyromagnetic ratio γ were calculated from the resonance fields in perpendicular and parallel orientations. The cubic anisotropy field $\mathbf{H}_{k1}$ was determined from the azimuth dependence of the resonance field,[8] and the Gilbert magnetic damping parameter α was determined from the half-width of the absorption line ΔH, $\alpha=\Delta H\gamma/\omega$, where ω is the angular frequency of the microwave field.

The SWR spectra were recorded on an RÉ-1301 rf spectrometer (frequency of the microwave field 9.34×10^9 Hz) at room temperature. The magnetic field was measured by nuclear magnetic resonance using a magnetometer. Since the SWR spectra were recorded at a constant frequency ω, the dispersion curves were plotted as the difference between the resonance fields of the zeroth and nth modes (H_0-H_n) ver-

TABLE I. Parameters of the investigated films.

Sample No.	Layer No.	Composition	h, μm	$4\pi M$, G	α	H_k^{eff}, Oe	γ, 10^7 Oe^{-1}s^{-1}
1	1	$(BiSmTm)_3(FeGa)_5O_{12}$	1.1	600	0.09	2280	1.75
	2	$(GdYTm)_3(FeGa)_5O_{12}$	0.4	600	0.015	480	1.57
2	1	$(YSmLuCa)_3(FeGe)_5O_{12}$	2.1	470	0.15	1210	1.74
	2	$Y_{2.98}Sm_{0.02}Fe_5O_{12}$	0.56	1740	0.003	−1715	1.76
3	1	$Sm_{1.2}Lu_{1.8}Fe_5O_{12}$	1.4	1760	0.12	790	1.76
	2	$Y_{2.98}Sm_{0.02}Fe_5O_{12}$	0.54	1740	0.003	−1715	1.76
4	1	$(SmEr)_3Fe_5O_{12}$	1.8	1330	0.2	96	1.38
	2	$Y_{2.98}Sm_{0.02}Fe_5O_{12}$	0.84	1740	0.003	−1715	1.76
5	1	$(YSmLuCa)_3(FeGe)_5O_{12}$	2.1	470	0.15	1210	1.74
	2	$Y_{2.98}Sm_{0.02}Fe_5O_{12}$	0.56	1740	0.003	−1715	1.76
	3	$(YSmLuEr)_3(FeGa)_5O_{12}$	1.3	830	0.16	190	1.51

sus $(n+1/2)^2$ for two-layer films and versus $(2n+1)^2$ for three-layer films. Here $n=0,1,2,\ldots$ enumerates the spin-wave mode orders. The parameters of the investigated films are summarized in Table I. Note that the damping parameters in one or two outside layers of the film (pinning layers) increases in order of increasing sample number. As a result, the spin pinning mechanism changes from dynamical[5] to dissipative.[9]

We infer from the results that the angular dependence of the SWR spectrum depends strongly on the values of α in the pinning layer. If the values of α in both layers are sufficiently small (sample 1), the SWR spectrum, as in Refs. 5 and 6, undergoes significant changes as the angle Θ_H is varied. The form of the spectrum, the number of modes, and the dispersion curves differ considerably between perpendicular and parallel orientations (Figs. 1a–1c). For $\Theta_H \cong 65°$ only one peak is observed in the spectrum, and in the interval $\Theta_H \cong 65-90°$ the resonance absorption spectrum consists of two peaks corresponding to the zeroth modes of each layer. A difference of more than an order of magnitude between the values of α in the layers (sample 2) produces a different angular dependence of the SWR spectrum. In perpendicular and parallel orientations the spectra become similar, and the slopes of the dispersion curves differ insignificantly (Fig. 2a). At an angle $\Theta_H \cong 40-50°$, however, only one absorption peak is again observed in the spectrum. In sample 3, where the value of α in one of the layer is raised even higher, even though the number of excited modes at intermediate angles Θ_H ($\Theta_H = 35-45°$) decreases, there is no longer a transformation of the entire spectrum into a lone zeroth mode from both layers (Fig. 3). The dispersion curves for perpendicular and parallel orientations are close, but as in sample 2, the number of excited spin-wave modes in the perpendicular orientation is somewhat higher than in the parallel orientation. When the value of α in the pinning layer increases to 0.2–0.85, the SWR spectrum becomes essentially isotropic (Fig. 4). At intermediate angles the slopes of the dispersion curves increase somewhat, owing to the influence of a variation in the equilibrium orientation of the magnetization as the SWR spectrum is being recorded.[10]

One conspicuous feature is the fundamentally different character of the angular dependences of the SWR spectra of three-layer films as opposed to two-layer films. Whereas in

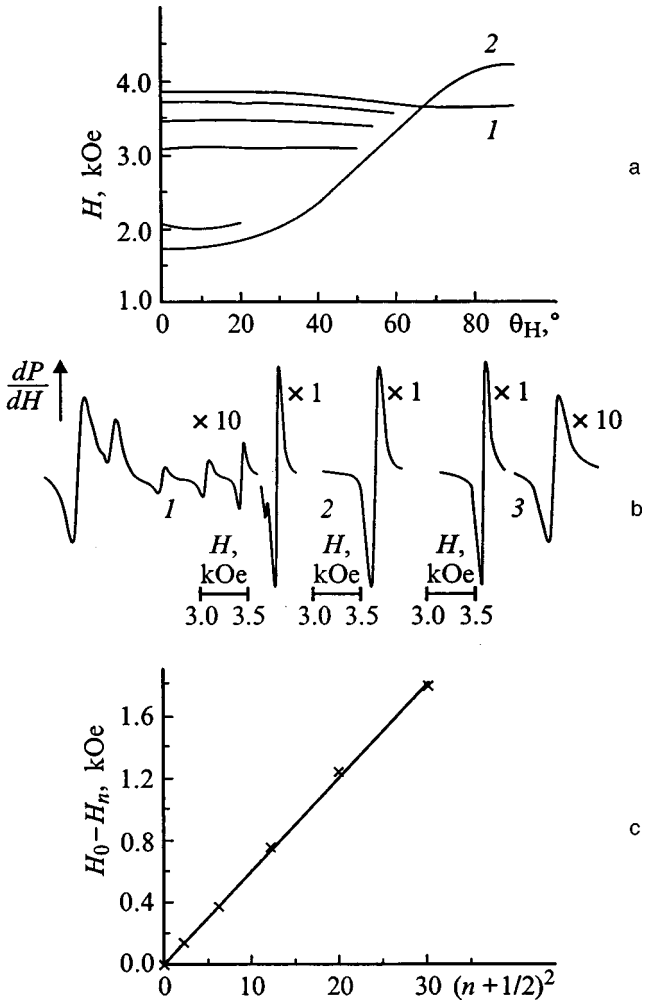

FIG. 1. a: Angular dependences of the uniform resonance fields in the excitation (*1*) and pinning (*2*) layers (the other curves represent the angular dependences of the resonance fields of excited spin-wave modes); b: SWR spectra for: $\Theta_H=0$ (*1*); $0<\Theta_H<90°$ (*2*); $\Theta_H=90°$ (*3*); c: dependence of H_0-H_n on $(n+1/2)^2$ for perpendicular (×) orientation of sample 1.

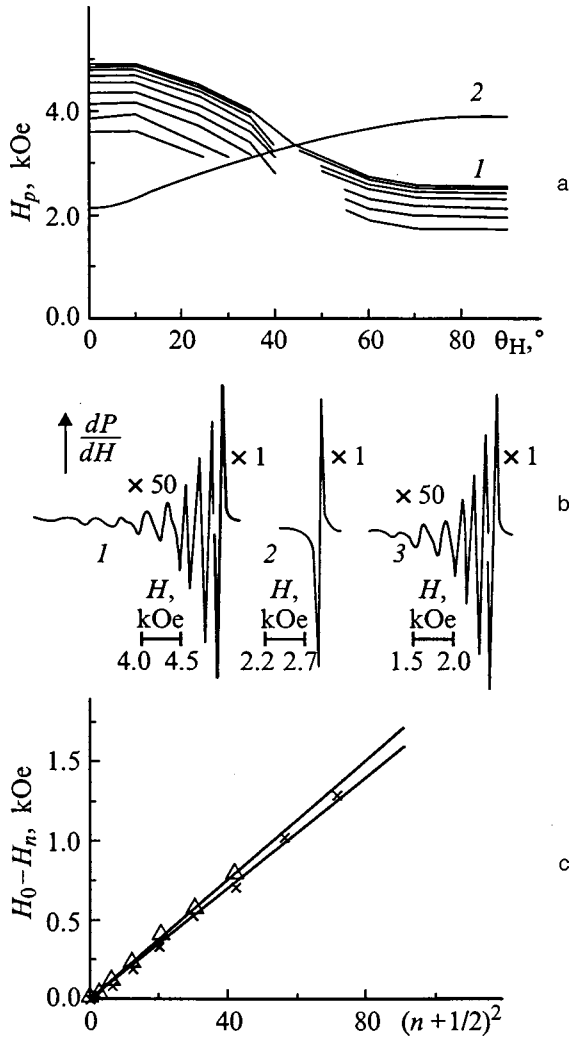

FIG. 2. a,b: The same as Fig. 1, for sample 2; c: dependences of H_0-H_n on $(n+1/2)^2$ for perpendicular (×) and parallel (△) orientations.

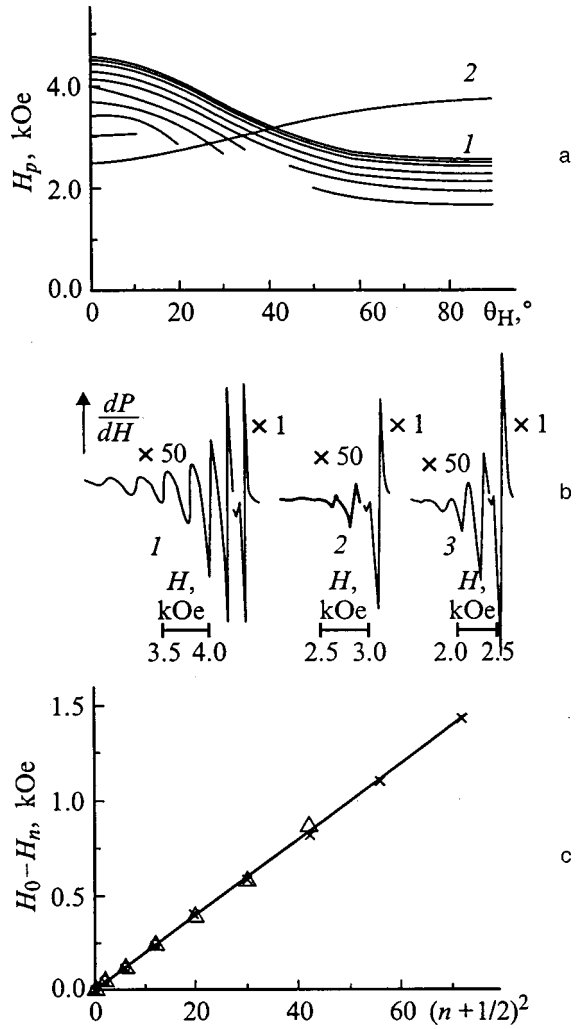

FIG. 3. The same as Fig. 2, for sample 3.

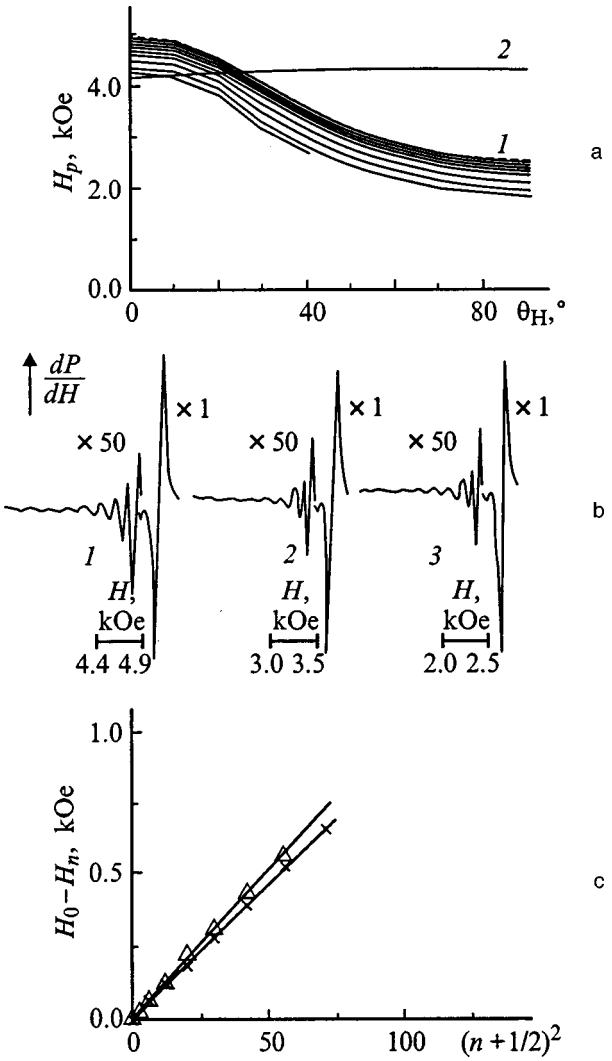

FIG. 4. The same as Fig. 2, for sample 4.

two-layer films the number of peaks decreases and then increases as Θ_H is varied from 0 to 90°, in a three-layer film (sample 5) this process is observed twice (Fig. 5). Moreover, in two-layer films, beginning with a high mode order, the excitation of successive spin-wave modes ceases, as Θ_H is varied, while in three-layer films, at approximately $\Theta_H = 12°$, only odd-order (intermediate) peaks vanish. Then in the interval of angles $40 \leqslant \Theta_H \leqslant 55°$ they reappear. Their intensity gradually increases and attains a maximum, becoming comparable with the intensity of the previously observed peaks. A further increase in the angle once again leads to the reverse process. The variation of the SWR spectrum of a three-layer film is accompanied by a clearly pronounced modification of the dispersion curve (Fig. 5c). In contrast with two-layer films (Figs. 1c–4c), the slope of the dispersion curve $H_0 - H_n = f(2n+1)^2$ at intermediate angles decreases considerably and then at $\Theta_H = 90°$ comes back to the original value. In three-layer films for which both pinning layers have a high value of α ($\alpha \geqslant 0.24$), the number of modes, as in the two-layer sample 4, does not depend on the angle Θ_H.

The multitude of angular dependences of the SWR spectra can be explained as follows. In films having a low damping parameter the principal spin pinning mechanism in both layers is dynamical. It follows from earlier results[5,6] that the localized modes excited most intensely by a homogeneous microwave field are those which are harmonic in one layer (the layer with the stronger field for uniform resonance H_{01}) and decay exponentially in the other layer, which in external fields stronger than its uniform resonance field H_{02} is a reactive (elastic) medium for spin waves[10] and thus induces spin pinning. For fields H smaller than H_{02} and H_{01} both layers become dispersive media for spin waves, and the efficiency of excitation of spin-wave resonance drops abruptly as a result. In this case spins are not pinned either at the free surfaces or at the interface between layers, and there is no longer any reason for the spectrum to exhibit the sharply defined discreteness observed for $H_{01} > H > H_{02}$. For a dynamical pinning mechanism, therefore, the interval of fields in which spin waves are strongly excited is bounded by H_{01} and H_{02}. Inasmuch as the interval $H_{01} - H_{02}$ decreases as Θ_H is varied, this behavior is accompanied by a decrease in the number of excited modes. At a certain angle Θ_H the fields H_{01} and H_{02} become equal, causing the spectrum to lose all modes except the one common zeroth mode. Finally,

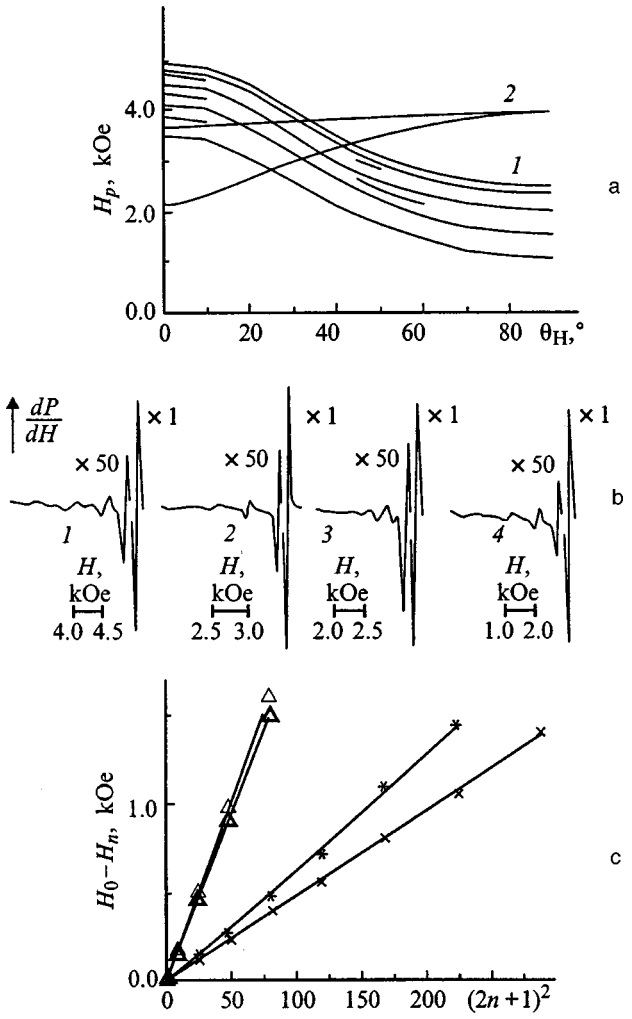

FIG. 5. a: The same as Fig. 1, for sample 5; b,c: $\Theta_H=0$ (×, *1*); 30° (▲, *2*); 45° (*, *3*); 90° (△, *4*).

with a further increase in Θ_H the field H_{02} becomes larger than H_{01}, so that the second layer emerges as the region of excitation of localized harmonic modes. If the thickness, the exchange constant A, and the magnetization differ in the first and second layers, the result is a difference in the slopes of the dispersion curves

$$H_0-H_n=\frac{2A}{M}\frac{\pi^2}{h^2}\left(n+\frac{1}{2}\right)^2$$

for perpendicular and parallel orientations. This accounts for the angular dependences of the SWR spectra of two-layer films with small and close values of α in the layers.

When α is increased in one layer of a two-layer film or in the two outer layers of a three-layer film, the influence of the dissipative spin pinning mechanism begins to increase.[9] This mechanism essentially entails the onset of a spin-wave node at or near the interface of two exchange-coupled layers with sharply different values of the damping parameter. We note that the efficiency of the dissipative pinning mechanism depends not only on α, but also on A, M, and other parameters. Among the characteristic attributes of this mechanism is the isotropy of its action due to the isotropy of α. It follows from the experimental results that any possible anisotropy of α in the investigated iron garnet films does not exceed the error of measurement (6%) of this parameter. An increase in the value of α in one of the layers (the pinning layer) has the effect that the influence of the dissipative mechanism becomes stronger, and for both perpendicular and parallel orientation the layer having the smaller value of α (the excitation layer) persists as the region of excitation of harmonic spin-wave modes. However, as the uniform resonance fields in the layers merge (this happens, for example, in the case of sample 2 at $\Theta_H\cong 40-50°$), again only one mode is observed in the SWR spectra (Fig. 2). The invariance of localization of the region of excitation of spin-wave modes is evinced by the very slight variation of the slope of the dispersion curve as Θ_H is varied from 0 to 90°. An analysis of the evolution of the SWR spectrum by etch peeling further corroborates this result. The etching of a film whose top layer has the higher value of α reveals that no appreciable changes in the spectrum occur down to a certain thickness of the top layer, after which the zeroth modes vanish. But the etching of a film whose top layer has the smaller value of α discloses a smooth increase in the slope of the dispersion curve for both orientations, indicating a decrease in the thickness of the layer in which harmonic spin-wave modes are excited.

Further strengthening of the dissipative mechanism (sample 3) eliminates the transformation of the entire spectrum into a single common mode from both layers (Figs. 3a and 3b), even when H_{01} and H_{02} are equal. The greater number of spin-wave modes excited in perpendicular orientation relative to parallel orientation can be attributed to the fact that the pinning layer for spin waves excited by harmonic standing waves localized in the layer with the smaller parameter α is a reactive (elastic) medium in this case. The simultaneous action of dissipative and reactive properties leads to more rapid spin-wave damping and, hence, a higher degree of spin pinning. In parallel orientation, on the other hand, it follows from Figs. 2 and 3 that the pinning layer is a dispersive medium.

Large values of α ($\alpha\geqslant 0.2$) and $4\pi M$ in the pinning layer make the dissipative pinning mechanism dominant over the dynamical mechanism and, as a result, render the SWR spectrum essentially isotropic (Fig. 4).

In our opinion, the observed difference in the dependences of the spectra of two-layer and three-layer films on $\Theta_H=0$ can account for the transitions in a three-layer film from asymmetric to symmetric boundary conditions, and vice versa, as Θ_H is varied. For symmetric boundary conditions (spin pinning at both boundaries of the excitation layer) modes that span the thickness of the excitation layer with an odd number of half wavelengths are excited by a homogeneous microwave field. Asymmetric boundary conditions (pinning at only one boundary of the excitation layer or nonequivalence of the degrees of pinning at the boundaries of the excitation layer) can be exploited to generate modes having an odd number of quarter wavelengths. Consequently, all other conditions being equal, approximately twice as many spin-wave modes are excited in the second case in the same interval of wave numbers. The latter consideration accounts

for the transformation of the spectrum of a three-layer film as Θ_H is varied. This transformation can take place by several different scenarios. If the dissipative pinning mechanism is dominant at both boundaries of the excitation layer, the spectrum remains unchanged, as in the two-layer sample 4. This result has been observed in our experiments. But when the dissipative pinning mechanism prevails at one boundary, and the dynamical mechanism is dominant at the other boundary, intermediate modes begin to be excited in the spectrum as the uniform resonance fields of the excitation layer and the pinning layer with a relatively small parameter α. This happens in the interval of angles Θ_H where the indicated uniform resonance fields are close to each other. An increase in Θ_H has the effect of increasing the difference between H_{0i} in the layers, strengthening the degree of pinning, and restoring the former pattern of the spectrum, i.e., the spectrum corresponding to symmetric boundary conditions. This mechanism also explains the most striking result, the fact that in the SWR spectrum of a three-layer film (Fig. 5) the number of modes decreases and increases twice in the interval of angles $\Theta_H(0-90°)$. In sample 5 both pinning layers have relatively small values of α and $4\pi M$, but their gyromagnetic ratios and effective anisotropy fields differ, thereby creating a difference in the angular dependences of the uniform resonance fields. Consequently, as Θ_H is varied, the uniform resonance field in the excitation layer approaches H_0, first in one pinning layer and then in the other. This event diminishes the degree of spin pinning, first at one boundary and then at the other boundary of the excitation layer. Consequently, as Θ_H is varied from 0 to 90°, there are two transitions from asymmetric to symmetric boundary conditions and vice versa, and this process, in turn, alters the number of absorption peaks in the spectrum. We note that only higher-order intermediate modes having the highest sensitivity to the degree of pinning have been observed in a number of cases. Obviously, the angular dependences of the SWR spectra observed in a three-layer film with identical magnetic parameters in the outer layers should be similar to those of two-layer films with corresponding pinning layers. The only difference is an approximately twofold reduction in the number of excited spin-wave modes.

To summarize, we have shown that different scenarios are possible for the transformation of the SWR spectrum of a multilayer film as the angle between the field **H** and the film is varied, depending on the symmetry of the boundary conditions and the characteristics of the layers of the film.

This work has received support from the Russian Research Foundation, Grant No. 98-02-03320.

[1] N. M. Salanskiĭ and M. Sh. Erukhimov, *Physical Properties and Applications of Magnetic Films* [in Russian] (Nauka, Novosibirsk, 1975).
[2] S. L. Vysotskiĭ, G. T. Kazakov, M. L. Kats, and Yu. A. Filimonov, Fiz. Tverd. Tela (St. Petersburg) **35**, 1190 (1993) [Phys. Solid State **35**, 665 (1993)].
[3] B. Hillebrands, Phys. Rev. B **1**, 530 (1990).
[4] C. Wilts and S. Prasad, IEEE Trans. Magn. **MAG-17**, 2405 (1981).
[5] B. Hoekstra, R. P. van Stapele, and J. M. Robertson, J. Appl. Phys. **48**, 382 (1977).
[6] A. M. Grishin, V. S. Dellalov, V. F. Shkar *et al.*, Phys. Lett. A **140**, 133 (1989).
[7] A. M. Zyuzin, V. N. Van'kov, and V. V. Radaĭkin, Pis'ma Zh. Tekh. Fiz. **17**(23), 65 (1991) [Sov. Tech. Phys. Lett. **17**, 848 (1991)].
[8] A. M. Zyuzin, V. V. Radaĭkin, and A. G. Bazhanov, Zh. Tekh. Fiz. **67**(2), 35 (1997) [Tech. Phys. **42**, 155 (1997)].
[9] A. M. Zyuzin, N. N. Kudel'kin, V. V. Randoshkin, and R. V. Telesnin, Pis'ma Zh. Tekh. Fiz. **9**, 177 (1983) [Sov. Tech. Phys. Lett. **9**, 78 (1983)].
[10] F. S. Crawford, Jr., *Waves* [McGraw-Hill, New York, 1968; Nauka, Moscow, 1974, 527 pp.].

Translated by James S. Wood

Relativistic backward wave oscillator using a selective mode converter

É. B. Abubakirov, A. N. Denisenko, N. F. Kovalev, E. A. Kopelovich, A. V. Savel'ev, E. I. Soluyanov, M. I. Fuks, and V. V. Yastrebov

Institute of Applied Physics, Russian Academy of Sciences, 603600 Nizhny Novgorod, Russia
(Submitted June 4, 1998)
Zh. Tekh. Fiz. **69**, 102–105 (November 1999)

A new version of the relativistic backward wave oscillator (BWO) is proposed and investigated experimentally, where the cutoff (for the working mode mode) taper at the cathode end is replaced by a selective Bragg-type mode converter. In the experimental BWO model, which operates in the three-centimeter range and is equipped with a mode converter based on a slightly corrugated waveguide, a radiated power of 700 MW in pulses of duration up to 100 ns with an output spatial structure similar to a Gaussian wave beam is obtained at an accelerating potential of 0.8 MV and a focusing magnetic field of 7 kOe. © *1999 American Institute of Physics.* [S1063-7842(99)01811-5]

INTRODUCTION

In the conventional structure of a high-power relativistic backward wave oscillator (BWO) the output of microwave radiation from the interaction space takes place through a collector region at zero potential. This condition is usually achieved by means of a cutoff taper, which reflects the working mode of the BWO from the cathode end of the electrodynamic system without disrupting transmission of the electron beam. This device configuration permits oversize elements to be used in the rf channel; such elements — diffractional output and large-aperture conical-horn antenna in particular — are best suited for high-power microwave radiation. However, the application of a cutoff taper imposes a number of restrictions on the output characteristics of the BWO. First, because the cutoff cross section functions simultaneously as a wave reflector and a diaphragm for the electron beam, one of the higher modes is generally used as the working mode in a BWO with an oversize electrodynamic system. The radiation therefore has a complex distribution, whereas a simple structure such as a wave beam is most often required for practical use. Second, the reflection of the working mode leads to amplification of the rf electric field in the interaction space of the BWO, increasing the risk of rf breakdown (as a rule, all electrodynamic systems in high-power relativistic BWOs show traces of erosion at antinodes of the electric field). Moreover, spurious interaction of the electron beam with a reflected TM mode can diminish the efficiency and stability of operation of the BWO and distort the structure of the output radiation.[1,2]

BACKWARD WAVE OSCILLATOR WITH A BUILT-IN MODE CONVERTER

The above-mentioned unfavorable factors affecting the structure of the output radiation can be alleviated to a large extent by replacing the cutoff taper with a selective mode converter, which can be used to change the working mode of the BWO into a copropagating wave of any desired structure. Simple radiation structures are the most attractive for practical applications, and it is preferable, therefore, to use one of the lower modes as the copropagating wave transporting radiation from the interaction space. Inasmuch as lower mode orders have a higher group velocity and, hence, a lower electric field amplitude, the risk of rf breakdown is diminished. When a converter having both frequency and mode selectivity is used, the degree of coherence of the BWO output radiation is improved. The output of microwave power from a low-mode interaction space facilitates the solution of another practical problem of utmost importance in the operation of devices using a high pulse repetition rate: how to raise the efficiency of the BWO in the presence of a weak focusing magnetic field.[3,4]

Finally, if the working mode of the BWO is converted into a copropagating hybrid mode of a slightly corrugated waveguide, with a structure similar to that of a TE mode of a regular cylindrical waveguide, the vanishingly small longitudinal component of the electric field reduces spurious interaction of the electron beam with this codirectional wave to a negligible level.

SELECTIVE MODE CONVERTER FOR A BACKWARD WAVE OSCILLATOR

The most practical mode converter configuration for a BWO is a Bragg reflector in the form of a cylindrical waveguide segment with a corrugated wall.[5] This type of converter is compatible with a high-power electron beam and has a sufficiently high electrical resistance. The operation of the Bragg reflector is based on resonance scattering of the working mode by the periodic corrugation. In a cylindrical waveguide with a periodic n-start helical or axisymmetrical $(n=0)$ corrugation the azimuth orders of the coupled modes are given by the relations

$$m_{i1} \pm n = m_{i2}, \tag{1}$$

and the sum of the longitudinal propagation constants of these modes is equal to the constant $\overline{h}_r = 2\pi/d_r$ of the periodic (with period d_r) system of the reflector:

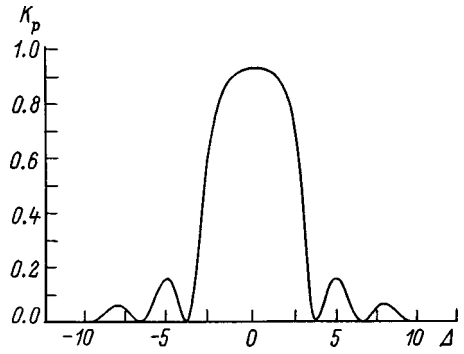

FIG. 1. Dependence of the mode conversion efficiency (power ratio) of the Bragg reflector on the detuning.

$$h_{i1}+h_{i2}=\overline{h_r} \tag{2}$$

(the index i enumerates the pairs of coupled modes).

Figure 1 shows the dependence of the mode (power) conversion efficiency K_P on the detuning

$$\Delta=(\overline{h_r}-h_{i1}-h_{i2})L/2 \tag{3}$$

for a reflector of length L (Refs. 5 and 6). To ensure reradiation of the working mode of the BWO into the desired copropagating mode and, at the same time, eliminate reradiation into spurious modes, the mode conversion band for the working field structure (Fig. 1) must not overlap the bands of adjacent structures, whose spectrum

$$k_i=\frac{1}{2\overline{h_r}}[(\overline{h_r^2}+k_{i1}^2+k_{i2}^2)^2-4k_{i1}^2k_{i2}^2]^{1/2}/2\overline{h_r} \tag{4}$$

($k_{i1}=2\pi/\lambda_i$, λ_i is the wavelength, and k_{i1} and k_{i2} are the transverse wave numbers) is given by Eq. (2). The frequency difference between the working and adjacent resonance structures is a maximum when the working structure involves the rf field of a quasi-transverse mode having a frequency close to the cutoff frequency. This is the situation in the Bragg reflector used in BWOs.

BACKWARD WAVE OSCILLATORS UTILIZING CYCLOTRON-RESONANCE MODE SELECTION

For the practical implementation of an oscillator with a built-in mode converter, we have chosen a BWO operating in the TM_{02} mode. Single-mode oscillation in the multimode interaction space of a BWO operating under the conditions of Čerenkov synchronism with the -1st spatial harmonic of the working mode,

$$\omega=(\overline{h})\nu, \tag{5}$$

is achieved by mode selection based on the resonance interaction of a fast cyclotron wave of an electron beam guided by a longitudinal magnetic field H with the fundamental of the working mode,[7] i.e., under the condition

$$\omega+h\nu=\omega_H. \tag{6}$$

Here ω and h are the frequency and longitudinal wave number of the working mode of the BWO, ν is the translational electron velocity, $\overline{h}=2\pi/d$, d is the period of the electrodynamic system of the BWO, $\omega_H=eH/mc\gamma$ is the cyclotron frequency, e and m are the charge and rest mass of the electron, c is the speed of light, and γ is the relativity factor.

If electrons entering the interaction space do not have transverse velocities and if conditions (5) and (6) hold simultaneously, i.e., if the magnetic field is chosen to satisfy the relation

$$\omega_H=\overline{h}, \tag{7}$$

the electromagnetic energy of the fast electron-beam cyclotron wave is absorbed.[8] It follows from condition (7), which does not involve any individual characteristics of the modes, that oscillation is suppressed for all modes simultaneously. One exception is a mode for which the intensity of cyclotron interaction with the rotating transverse structure of the field, being proportional to the factor $J_{m\pm1}^2(\nu_{mn}r/R)$, is equal to zero. Here r is the radius of the annular electron beam, R is the average radius of the electrodynamic system, ν_{mn} is the nth root of the Bessel function $J_m(\nu)=0$ for the TM_{mn} mode or the derivative of the Bessel function $J_m'(\nu)=0$ for the TE mn mode, and the signs $+$, $-$ refer to opposite directions of rotation of the transverse wave structure. For a TM_{02} working mode cyclotron absorption is absent for $r/R=0.693$. It is important to note that the strong focusing magnetic field required for the efficient operation of Čerenkov oscillators in the centimeter wavelength range is greatly diminished when the cyclotron-resonance technique is used for mode selection.

The operation of centimeter-range relativistic BWOs of standard configuration are not very efficient for magnetic fields smaller than the resonance value (7), because the range of weak magnetic fields contains another cyclotron resonance, which results in the strong absorption of rf power. This resonance takes place with the forward wave

$$\omega-h\nu=\omega_H, \tag{8}$$

which transports radiation from the interaction space of the BWO. The interval of magnetic fields H between the resonances (6) and (8),

$$e(H_2-H_1)/mc^2=2h(\gamma^2-1)^{1/2}, \tag{9}$$

in which the oscillator is capable of operating, is fairly narrow, because the propagation constant of the working mode h in a BWO is usually much smaller than k, and the low operating efficiency of the device in this interval is attributable to the influence of nearby cyclotron resonances.

By using a lower mode to transport radiation from the interaction space it is possible to significantly extend this interval of magnetic fields H to

$$e(H_2-H_1)/mc^2\approx(k+h)(\gamma^2-1)^{1/2} \tag{10}$$

as a result of the shift of the resonance (8) with the copropagating wave, for which the phase velocity is close to the speed of light, toward lower values of H. This effect permits the magnetic field to be chosen so as to essentially reduce to zero the influence of cyclotron effects on the efficiency of the induced Čerenkov radiation.

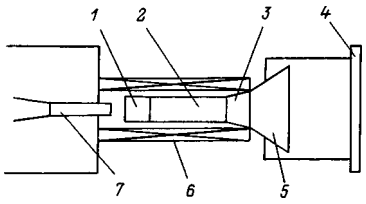

FIG. 2. Schematic view of the investigated oscillator.

EXPERIMENTAL INVESTIGATION OF A BACKWARD WAVE OSCILLATOR UTILIZING A BRAGG REFLECTOR

Experiments have been carried out with the oscillator shown schematically in Fig. 2. The main calculated parameters of the generator are as follows: electron energy 0.8 MeV; beam current 6 kA; the electrodynamic system *2* of the BWO comprises a cylindrical waveguide segment with a periodic, axisymmetrical, sinusoidal corrugated profile; the oscillation frequency in the $TM_{0.2}$ mode is 9.1 GHz, and the expected radiation power is greater than 0.5 GW. The electrodynamic system *1* of the Bragg reflector comprises a cylindrical waveguide segment with a single-start ($n=1$) helical, sinusoidal corrugated profile; the calculated efficiency of conversion of the power of the incident TM_{02} mode into the power of a reflected rotating TE_{11} mode at a frequency of 9.1 GHz is $K_P=91.2\%$, and the conversion bandwidth is 4%.

A high-current electronic amplifier with an explosive-emission cold cathode *7* was used in the experiments. The accelerating potential on the magnetically insulated diode was varied in the interval 0.7–1.1 MV in a pulse of duration 200 ns, the beam current ranged from 5 to 7 kA, and the beam diameter was varied from 41 to 47 mm. A solenoid *6* was used to focus the electron beam, producing a magnetic field up to 20 kOe in 10-ms pulses. Radiation was extracted from the electrodynamic system *2* of the oscillator by a horn antenna *3*, which generated a wave beam, through a resonant vacuum-tight polyethylene window *4* of diameter 40 mm, whose thickness was made equal to the wavelength to ensure mechanical strength. The fully developed electron beam impinged on a collector *5* situated outside the rf field. In this collector the filling of the rf channel with plasma generated by electron impingement was much slower than in the collector serving as part of the rf section.[9]

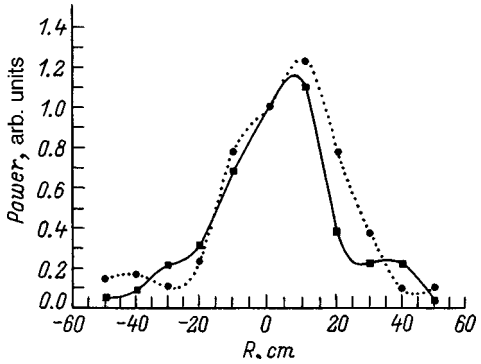

FIG. 4. Intensity distribution of radiation from the BWO at a distance of 4 m from the radiating cone.

The frequency of the generated pulse was measured at the center of the beam pattern by means of a tunable guided-wave band filter. The measurements showed that the oscillation frequency was close to the calculated value.

The polarization of the output radiation was determined by means of a device utilizing a dielectric plate situated in the diametrical plane of the cylindrical waveguide. The dimensions of the plate were chosen so that, at the working frequency, circularly polarized radiation incident on the cylindrical waveguide would be converted into linear polarized radiation and transfer into a rectangular waveguide smoothly joined to it when the plate was mounted at a 45° angle relative to the rectangular waveguide. Accordingly, at an angle of 135° radiation is reflected from the rectangular waveguide. Figure 3 shows BWO signals after such a polarizer for two positions of the dielectric plate, showing that the polarization of the observed radiation is close to circular.

The spatial structure of the output radiation was measured for vertical and horizontal polarizations in the plane at a distance of 4 m from the radiating cone *3* in the horizontal and vertical directions. The measurement results (Fig. 4) are in good agreement with the standard distribution obtained in excitation of the output section of the BWO by the TE_{11} mode of the cylindrical waveguide, which was then converted into a wave beam in the radiating cone. Consequently, measurements of the frequency, polarization, and structure of the output radiation confirm that the BWO operates, as expected, in the TM_{02} mode, and the reflector band overlaps with the emission band of the BWO.

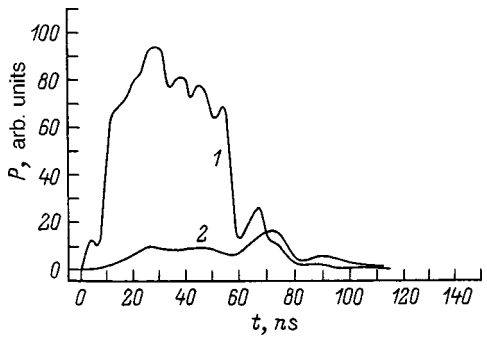

FIG. 3. Oscillograms of the envelopes of microwave pulses at the output of a polarizer adjusted for left- (*1*) and right- (*2*) circularly polarized radiation.

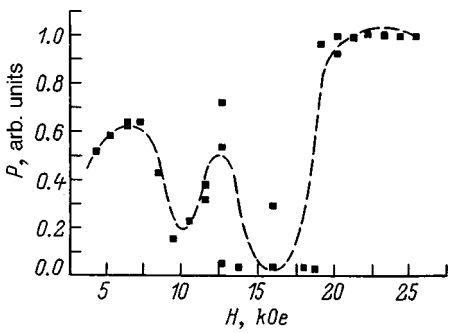

FIG. 5. Dependence of the oscillator output power on the focusing magnetic field.

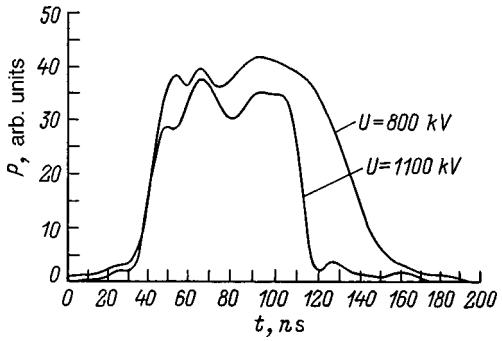

FIG. 6. Oscillograms of the envelopes of the output radiation for different accelerating potentials.

Of special interest for this version of BWO is the dependence of the output power on the guiding magnetic field (Fig. 5). Since the radius of the electron beam in the BWO is made equal to $r/R=0.693$ to minimize the intensity of cyclotron interaction of electrons with the working TM_{02} mode, the zone of cyclotron absorption is relatively small at $H=10$ kOe. In the vicinity of $H=16$ kOe, however, a second cyclotron absorption zone emerges, corresponding to resonance ($\omega_H=\overline{h}_r v$) reradiation in the working volume of the Bragg reflector of a counterpropagating TM_{02} electromagnetic wave into fast electron-beam cyclotron waves. In addition, as expected, the use of waves having a relatively low phase velocity to tap energy from the BWO made it possible to sharply define the window of transparency in the range of weak magnetic fields. By optimizing the output power with respect to the radius of the electron beam at weak magnetic fields in the vicinity of $H=7$ kOe we have been able to attain 0.7 GW in 100-ns pulses at approximately 15% efficiency.

An increase in the accelerating potential (and, accordingly, the beam current) does not produce the expected increase in the output power. The radiation pulses become shorter and irregular (see Fig. 6), most likely indicating the onset of rf breakdown in the microwave section (in the electrodynamic system of the BWO, at the output window, or both).

CONCLUSION

The results of the experimental investigation have demonstrated a significant improvement in the performance characteristics of a relativistic backward wave oscillator when the conventional cutoff taper at the cathode end of the electrodynamic system is replaced by a selective Bragg reflector.

[1] É. B. Abubakirov and M. I. Petelin, Zh. Tekh. Fiz. **58**, 1085 (1988) [Sov. Phys. Tech. Phys. **33**, 635 (1988)].
[2] S. D. Korovin, M. I. Polevin, A. M. Roĭtman, and V. V. Rostov, Pis'ma Zh. Tekh. Fiz. **20**(1), 12 (1994) [Tech. Phys. Lett. **20**(1), 5 (1994)].
[3] E. B. Abubakirov, M. I. Fuchs, N. G. Kolganov *et al.*, in *Strong Microwaves in Plasmas*, Vol. 2 (IAP RAS, Nizhny Novgorod, 1996), pp. 810–828.
[4] D. Clunie, G. Mesyats, M. I. Osipov *et al.*, in *Strong Microwaves in Plasmas*, Vol. 2 (IAP RAS, Nizhny Novgorod, 1996), pp. 886–902.
[5] B. Z. Katsenelenbaum, *Theory of Irregular Waveguides with Slowly Varying Parameters* [in Russian] (Izd. AN SSSR, Moscow, 1981).
[6] N. F. Kovalev, I. M. Orlova, and M. I. Petelin, Izv. Vyssh. Uchebn. Zaved., Radiofiz. No. 11, 783 (1968).
[7] É. B. Abubakirov, V. I. Belousov, V. N. Varganov *et al.*, Pis'ma Zh. Tekh. Fiz. **9**, 533 (1983) [Sov. Tech. Phys. Lett. **9**, 230 (1983)].
[8] É. B. Abubakirov, N. S. Ginzburg, N. F. Kovalev, and M. I. Fuks, Radiotekh. Elektron. (Moscow) **34**, 1058 (1989).
[9] A. M. Bykov, V. P. Gubanov, A. V. Gunin *et al.*, in *Relativistic Radio-Frequency Electronics* [in Russian], No. 5 (IPF AN SSSR, Gorki, 1988), pp. 101–124.

Translated by James S. Wood

Equivalent parameters of a Josephson junction in a microwave SQUID structure

O. G. Vendik and I. S. Danilov

St. Petersburg State Electrical-Engineering University, 197376 St. Petersburg, Russia
(Submitted June 18, 1998)
Zh. Tekh. Fiz. **69**, 106–112 (November 1999)

The equivalent parameters of a Josephson junction in a microwave SQUID structure are calculated on the basis of relations obtained as a result of an analysis of the operation of an rf SQUID. This analysis is based on the sawtooth variation of the voltage on the resonator as a function of the constant flux bias. The quantitative characteristics permit regarding the Josephson junction as a linear impedance in the rf or microwave circuit, whose real and imaginary parts are controlled by the constant magnetic flux passing through the SQUID loop.
 [S1063-7842(99)01911-X]

INTRODUCTION

The operating principle of an rf SQUID has been the subject of many publications (see, for example, Refs. 1 and 2). The first experiments on the operation of a microwave SQUID were performed long ago.[3] Novel designs for microwave SQUIDS have been planned,[4,5] and a theory for the operation of a SQUID at 77 K has been put forward.[6] Questions concerning the quantum interference in a SQUID ring, the effect of thermal fluctuations on it,[1,7] etc. have been developed in fairly great detail. On the other hand, proper attention has not been focused on the calculation of microwave circuits. The microwave part of a SQUID is described in words as a high-Q resonator. The coupling of the resonator to external circuits is generally adjusted during an experiment and is not subjected to a detailed calculation.[5] Thus, the microwave elements of a SQUID are planned on the basis of approximate relations determined as a result of experimental work, i.e., there are no reliable computational formulas for developing and selecting the optimum operating regime of a microwave SQUID.

This paper describes a method for representing the equivalent parameters of a Josephson junction in an rf SQUID structure in the hysteretic operating regime, which permits optimization of the parameters of the microwave circuit.

The main element in the equivalent circuit of the SQUID is the impedance of the SQUID loop, which is inductively coupled to the resonator. In the hysteretic SQUID operating regime the time dependence of the amplitude of the voltage on the resonator corresponds to a relaxation process.[8] However, under the condition $T_P \ll \tau \ll T_M$, where T_P is the pumping period, τ is the relaxation time, and T_M is the measurement time, the output signal is a time-averaged quantity. Therefore, to find the response of the ring with the Josephson junction to a constant magnetic flux at $\tau \ll T_M$, the relaxation process can be ignored. The equivalent impedance of the SQUID loop is a function only of the constant external magnetic flux. Such a representation permits considerable simplification of the calculation of the equivalent circuit and can be used, in particular, to analyze and synthesize a microwave SQUID.

1. SPREAD OF THE SIGNAL CHARACTERISTIC

The equivalent circuit of a single-junction SQUID is shown in Fig. 1. According to experimental data, the plot of the voltage on the resonator V_T as a function of the constant bias magnetic flux $\Phi_{\rm DC}$ is a sawtooth characteristic.[3,5] This dependence is presented in Fig. 2 in normalized quantities:

$$W(x) = V_T(x)/V_{T,\max}, \tag{1}$$

where

$$V_{T,\max} = I_1 \omega L_T Q, \tag{2}$$

$x = 2\Phi_{\rm DC}/\Phi_0$, $Q = \omega L_T/R_T$ is the Q factor of the resonator without consideration of the losses introduced by the SQUID loop, ω is the pump frequency, and Φ_0 is the magnetic flux quantum.

We define the normalized spread of the signal characteristic as

$$\Delta W = \Delta V_T / V_{T,\max}, \tag{3}$$

where ΔV_T is the difference between voltages on the resonator for two different values of the bias flux: $\Phi_{\rm DC} = 0$ ($x = 0$) and $\Phi_{\rm DC} = \Phi_0/2$ ($x = 1$).

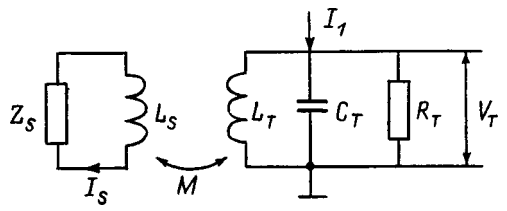

FIG. 1. Equivalent circuit of a single-junction SQUID: Z_S — equivalent impedance of the Josephson junction; L_S — inductance of the SQUID loop; M — mutual inductance; L_T — inductance of the resonator; R_T — active resistance of the resonator; C_T — capacitance of the resonator; I_1 — pump current; V_T — measured voltage.

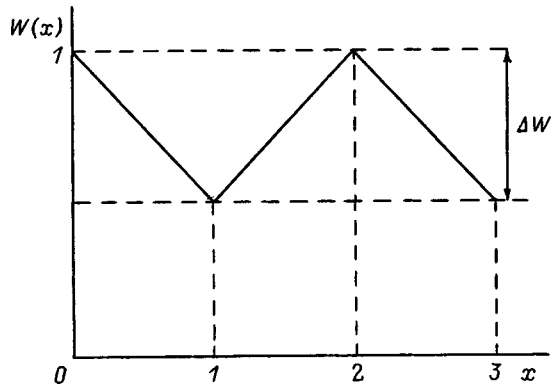

FIG. 2. Sawtooth variation of the normalized voltage on the resonator as a function of the normalized constant bias flux.

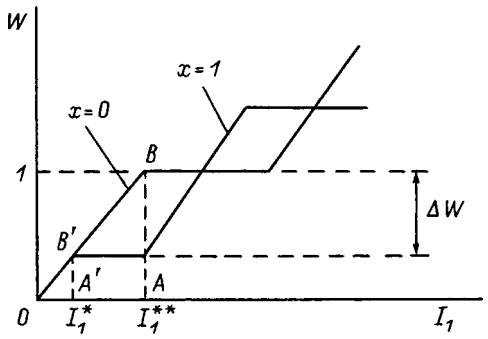

FIG. 3. Variation of the normalized voltage on the resonator as a function of the pump current for the case of coupling between a SQUID and a nearly optimal oscillator circuit.

Then the law governing the variation of the output signal as a function of the bias flux can be represented analytically as a periodic function of x, which, within the first half-period of the variation of W has the form

$$W(x)=(1-\Delta W\cdot x)\exp(i\psi(x)), \tag{4}$$

where $\psi(x)$ is the phase of the complex variable.

Typically $\Delta W\cong 0.5$–0.6. The sawtooth course of $|W(x)|$ is confirmed by reliable experimental data.[1–5] We do not know of any experimental or theoretical study of the variation of the phase of the voltage on the resonator $\psi(x)$ as a function of the constant bias flux.

When there is optimum coupling of the SQUID loop and the resonator,[1] $k^2Q_{\text{opt}}\approx 1$, where $k=M^2/(L_TL_S)$ is the coupling coefficient. The spread of the signal characteristic is greatest in this case. Figure 3 presents the output signal characteristic as a function of the pump current for a condition close to the optimum situation.

The working pump current I_1 is chosen in such a manner that at zero bias flux, $\Phi_{\text{DC}}=0$ $(x=0)$, the SQUID loop would have a purely imaginary impedance and would not absorb power from the resonator. When the bias flux $\Phi_{\text{DC}}=\Phi_0/2$ $(x=1)$, the dissipation of energy from the resonator is compensated during one rf oscillation period.[9] This corresponds to the current $I_1=I_1^{**}$ (Fig. 3). Thus, under the condition $x=0$ the impedance of the SQUID loop has only a reactive component, and when $x=1$, the impedance has both a reactive component and a significant active component.[1)]

It follows from consideration of the similar triangles AOB and $A'OB'$ in Fig. 3 that

$$\Delta W=1-I_1^*/I_1^{**}. \tag{5}$$

To find the dependence of ΔW on the fundamental SQUID parameter $l=2L_SI_C/\Phi_0$, where I_C is the critical current of the Josephson junction, the dependence of the total magnetic flux Φ on the external flux Φ_E must be analyzed. This dependence has the following form in quantities normalized to $\Phi_0/2\pi$:[9]

$$\varphi_E=\varphi+l\sin(\varphi), \tag{6}$$

where $\varphi_E=2\pi\Phi_e/\Phi_0$ and $\varphi=2\pi\Phi/\Phi_0$.

A plot of $\varphi(\varphi_E)$ for $l\approx 3$ is presented in Fig. 4. For the points of inflection (1), (2), (3), and (4) of the function $\varphi_E(\varphi)$ we can write $d\varphi_E/d\varphi=0$, which, with consideration of (6), gives

$$1+l\cos(\varphi)=0. \tag{7}$$

The flux at the points of inflection are functions of l. For a further analytical treatment it would be most convenient to work with the second point of inflection. For it we have

$$\varphi_2(l)=\pi-\arccos(1/l). \tag{8}$$

Using formula (6), for φ_{E2} we have

$$\varphi_{E2}(l)=\pi-\arccos(1/l)+(l^2-1)^{0.5}. \tag{9}$$

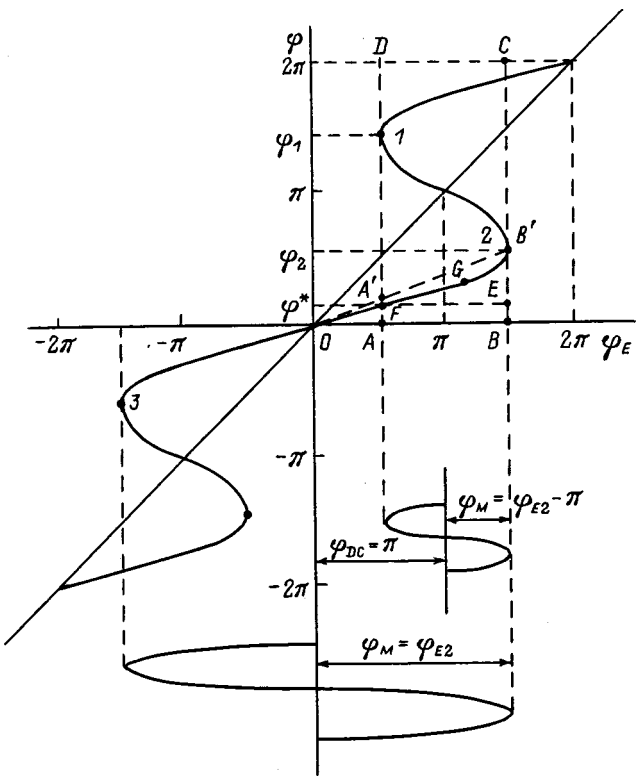

FIG. 4. Dependence of the normalized total magnetic flux φ through a SQUID loop on the normalized external flux φ_E.

Under the condition of the optimum choice of the working pump current (Fig. 4), the amplitude of the variation of the normalized rf magnetic flux imparted by the resonator to the SQUID loop will be

$$\varphi_{\mathrm{RF}}^{\max}=\varphi_{E2} \quad \text{for} \quad x=0,$$

$$\varphi_{\mathrm{RF}}^{\min}=\varphi_{E2}-\pi \quad \text{for} \quad x=1. \tag{10}$$

The coupling between the pump current and the normalized external flux can be expressed as

$$\varphi_E=\frac{2\pi}{\Phi_0}MQI_1. \tag{11}$$

Using (11), for the characteristic values of the pump current shown in Fig. 3 (I_1^* and I_1^{**}) we have

$$I_1^{**}=(\varphi_{\mathrm{RF}}^{\max}\cdot\Phi_0)/(2\pi MQ) \quad \text{for} \quad x=0,$$

$$I_1^{*}=(\varphi_{\mathrm{RF}}^{\min}\cdot\Phi_0)/(2\pi MQ) \quad \text{for} \quad x=1. \tag{12}$$

Substituting (12) into (5) with consideration of (10), we obtain

$$\Delta W(l)=\pi/\varphi_{E2}(l). \tag{13}$$

Table I presents values of the normalized spread of the signal characteristic, ΔW, for various values of the fundamental SQUID parameter l, according to calculations using formulas (9) and (13).

2. FORMULATION OF THE PROBLEM OF INTERPOLATING THE AVERAGED VALUE OF THE IMPEDANCE IMPARTED BY THE MICROWAVE SQUID LOOP TO THE RESONATOR

The law governing the variation of the rf voltage on the resonator represented by (4) is an experimental fact. Moreover, the value of ΔW was determined as a function of the fundamental SQUID parameter l (Table I). The use of the expression (4) permits avoidance of calculation of the relaxation of the voltage on the resonator and allows consideration of the averaged voltage on the resonator as the response to the averaged impedance imparted by the SQUID loop to the resonator.

Let us first consider the SQUID operating regime at the two points at $x=0$ and 1, at which there is no relaxation of the voltage on the resonator. The impedance values found at these points allow us to use (4) to interpolate the impedance dependence to all values of x in the range $0\leqslant x\leqslant 1$. Along the way it will be necessary to make some assumptions regarding the behavior of the phase of the voltage on the resonator as a function of x. In the end we should obtain the values of the complex impedance imparted by the SQUID loop to the resonator at all values of x.

Thus, the following sections are devoted to finding the impedance of a SQUID loop subjected to the effects of the optimum rf magnetic flux according to (7) and the two values of the constant bias flux $\Phi_{\mathrm{DC}}=0$ and $\Phi_{\mathrm{DC}}=\Phi_0/2$.

TABLE I. Values of the normalized spread of the output signal.

l	ΔW
1	1
1.5	0.92
2	0.82
2.5	0.73
3	0.66
3.5	0.6
4	0.55
4.5	0.51
5	0.47

3. VALUE OF THE SQUID LOOP IMPEDANCE FOR $\Phi_{\mathrm{DC}}=0$

To find the SQUID loop impedance for $\Phi_{\mathrm{DC}}=0$ we turn once again to the plot of $\varphi(\varphi_E)$ in Fig. 4. The loop is ''probed'' by a signal of the form

$$\varphi_{\mathrm{RF}}(t)=\varphi_M\cos(\omega t). \tag{14}$$

The flux through the loop $\varphi(t)$ is found as the solution of the nonlinear equation (6), which we rewrite as follows:

$$\varphi(t)+l\sin(\varphi(t))=\varphi_M\cos(\omega t). \tag{15}$$

Under the condition that

$$\varphi_M=\varphi_{E2}-\xi, \tag{16}$$

where $\xi\ll\varphi_{E2}$, the solution of (15) is obtained in the form of a single-valued function. The current in the SQUID loop is defined by the fundamental expression[1,2]

$$I_S(t)=I_C\sin(\varphi(t)). \tag{17}$$

We represent the current in the form of an expansion in Fourier series. We confine ourselves to the first term of the expansion

$$I_S(t)=A\cos(\omega t)+B\sin(\omega t), \tag{18}$$

where

$$A=\frac{1}{\pi}\int_{-\pi}^{\pi}I_S(t)\cos(t)\,dt, \tag{19}$$

$$B=\frac{1}{\pi}\int_{-\pi}^{\pi}I_S(t)\sin(t)\,dt. \tag{20}$$

Kirchhoff's equation for a ring has the form

$$L_\Sigma\frac{\partial I_S}{\partial t}+RI_S=\frac{\partial\Phi}{\partial t}, \tag{21}$$

where L_Σ is the inductance of the SQUID loop with allowance for the contribution of the inductance of the Josephson junction.

Allowing for the fact that $\Phi_M=\varphi_M(\Phi_0/2\pi)$ and taking into account the form of the probe signal (14), we can represent the expression (21) as the system of equations

$$-\omega L_\Sigma\cdot A+R\cdot B=-\omega\frac{\Phi_0}{2\pi}\varphi_M,$$

$$R\cdot A+\omega L_\Sigma\cdot B=0. \tag{22}$$

We introduce the concept of the ''normalized'' impedance of a Josephson junction

$$z_S=Z_S/\omega L_S, \tag{23}$$

TABLE II. Values of the reactive component of the impedance for $\Phi_{DC}=0$.

l	x_S
2	0.77
2.5	0.59
3	0.48
3.5	0.40
4	0.35
4.5	0.31
5	0.27

where $L_S=\Phi_0/(2\pi I_C)$.

Then, the normalized active and reactive components r_S and x_S have the form

$$r_S=R/\omega L_S, \quad x_S=(\omega L_\Sigma-\omega L_S)/\omega L_S. \tag{24}$$

Using the expressions (22) and (24), we can represent r_S and x_S in the form

$$r_S=-I_C\frac{\varphi_M}{l}\frac{B}{A^2+B^2}, \quad x_S=I_C\frac{\varphi_M}{l}\frac{A}{A^2+B^2}-1. \tag{25}$$

Under the condition assigned by (16), the active component of the impedance is equal to zero, i.e., $B=0$ and thus $r_S=0$. A calculation using formulas (25) gives the values of $x_S(l)$ for $\Phi_{DC}=0$, which are listed in Table II.

4. VALUE OF THE IMPEDANCE OF A SQUID LOOP AT A CONSTANT BIAS FLUX $\Phi_0/2$ ($x=1$)

a) *Finding the impedance using expansion in a Fourier series.* Let us consider the form of the time dependences of φ for the values of the amplitude of the radio-frequency signal assigned by the expression

$$\varphi_M=\varphi_{E2}-\pi\pm\xi. \tag{26}$$

When $\varphi_M=\varphi_{E2}-\pi-\xi$, the solution (15) is obtained in the form of a single-valued function, and the time dependence $\varphi(t)$ has the form shown in Fig. 5a. When $\varphi_M=\varphi_{E2}-\pi+\xi$, the form of the time dependence is determined by ''slippage'' of the magnetic flux quantum and has the form shown in Fig. 5b.

Taking into account that a phase change by 2π does not influence the representation of the signal in the form of an expansion in trigonometric functions, we can write the time dependence for the case (26) in the form of the dependence shown in Fig. 5c. In this case expansion into a Fourier series gives both the real and imaginary components. The results of the calculation are listed in Table III.

b) *Finding the active component using the expression for the area of the hysteresis loop.* The active component of the impedance of the SQUID loop can be obtained using the expression of the area of the hysteresis loop and the energy conservation law.[1,8] We again turn to Fig. 4, from which it is seen that to calculate the area of the hysteresis loop, twice the area of rectangle $ABEF$ and twice the area enclosed by $FEB'G$ must be subtracted from rectangle $ABCD$:

$$S(l)=ABCD-2ABEF-2FEB'G, \tag{27}$$

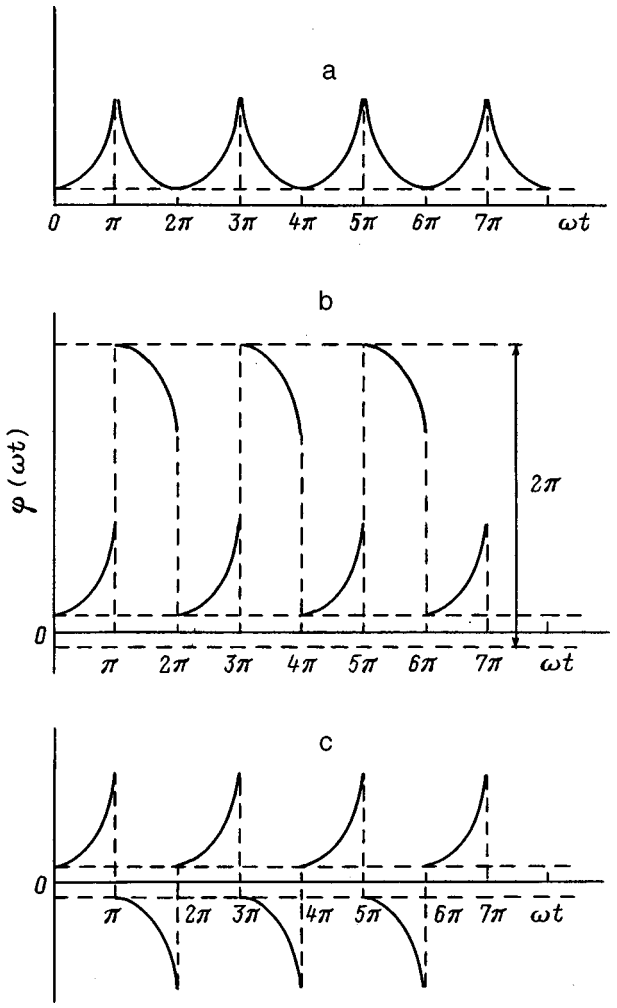

FIG. 5. Time dependences of the normalized total magnetic flux φ through a SQUID loop ($x=1$) corresponding to the conditions: a — $\varphi_M=\varphi_{E2}-\pi-\xi$, b — $\varphi_M=\varphi_{E2}-\pi+\xi$, c — $\varphi_M=\varphi_{E2}-\pi+\xi$, for $\omega t=[\pi/2,\pi]+\pi n$, where $n=0,1,2,\ldots$, i.e., the value of the normalized total magnetic flux φ has been reduced by 2π.

where

$$ABCD=2\pi(\varphi_{E2}(l)-\varphi_{E1}(l)), \tag{28}$$

$$2ABEF=2\left((\varphi_2(l)-\varphi^*(l))\varphi_{E2}(l)-\int_{\varphi^*(l)}^{\varphi_2(l)}(\varphi+l\sin(\varphi))\,d\varphi\right), \tag{29}$$

TABLE III. Values of the active and reactive components of the impedance for $\Phi_{DC}=\Phi_0/2$.

l	r_S	x_S
2	0.40	−0.98
2.5	0.48	−0.91
3	0.50	−0.84
3.5	0.50	−0.77
4	0.49	−0.72
4.5	0.47	−0.66
5	0.45	−0.62

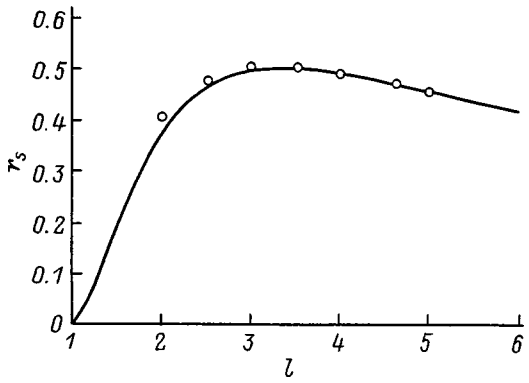

FIG. 6. Dependence of the active component of the equivalent impedance of the Josephson junction r_S. Circles — values calculated using the expansion of the current in a Fourier series; solid curve — values determined by calculating the area of the hysteresis loop.

$$2FEB'G=2(\varphi_{E2}(l)-\varphi_{E1}(l))\varphi^*(l). \tag{30}$$

In deriving the formula for the area of the hysteresis loop we used the condition (7) and the dependence assigned by (6). In this case it is possible to find exact analytical expressions for φ_{E1}, φ_{E2}, φ_1, and φ_2 as functions of l. An attempt to find an exact analytical expression for $\varphi^*(l)$ leads to cumbersome calculations. However, an approximate analytical function can be found from the relation for the similar triangles AOA' and BOB':

$$\varphi^*(l)=\varphi_2\varphi_{E1}/\varphi_{E2}. \tag{31}$$

It must be noted that the error in the estimate of the area of the hysteresis loop does not exceed 5%. For $l=5$ the value of $S(l)$ calculated from formula (27) agrees well with the data in Ref. 8.

For a bias flux equal to $\Phi_{DC}=\Phi_0/2$ ($x=1$) and the amplitude of the pump current corresponding to restoration of the energy in the resonator lost in the hysteresis cycle during one period, we have the energy conservation law

$$\frac{\omega}{2\pi}\frac{S(l)}{4\pi^2}\Phi_0^2\frac{1}{L_S}=\frac{1}{2}I_C^2R. \tag{32}$$

Formula (29) yields

$$R=\omega L_S r_S(l), \tag{33}$$

$$r_S(l)=S(l)/(\pi l^2). \tag{34}$$

A plot of $r_S(l)$ found using Eq. (34) and the numerical values of r_S in Table III, which were obtained as a result of expansion of the current into a Fourier series, are shown in Fig. 6.

5. DEPENDENCE OF THE SQUID LOOP IMPEDANCE ON BIAS FLUX

The circuit presented in Fig. 1 obeys the equality

$$\left(\frac{1}{i\omega L_T+(\omega M)^2/(i\omega L_S+Z_S)}+R_T^{-1}+i\omega C_T\right)V_T=I_1. \tag{35}$$

Utilizing, in particular, the smallness of the coupling coefficient k, we bring (35) into the form

$$\left(1-\omega^2L_TC_T+k^2\left(1-i\frac{Z_S}{\omega L_S}\right)^{-1}+i\frac{1}{Q}\right)\frac{V_T}{i\omega L_T}=I_1. \tag{36}$$

We introduce the notation

$$u=\frac{\omega_0^2-\omega^2}{\omega_0^2}Q. \tag{37}$$

The parameter u is the normalized detuning of the resonator. Using (37), we bring the expression (36) into the form

$$\left(iu-ip\left(1-i\frac{Z_S}{\omega L_S}\right)^{-1}+1\right)\frac{V_T}{\omega L_TQ}=I_1, \tag{38}$$

where

$$p=k^2Q. \tag{39}$$

Finally, for the normalized output voltage we have

$$W=\left(1+iu-i\frac{p}{1-iz_S}\right)^{-1}. \tag{40}$$

From (40) we find the normalized equivalent impedance of the SQUID loop

$$z_S=\frac{p}{-1-iu+1/W(x)}-i. \tag{41}$$

When $x=0$, the active component is equal to zero, i.e., $r_S=0$, and the normalized voltage equals unity, i.e., $W(0)=1$. Under this condition, we can use (40) to obtain the dependence for the detuning of the resonator which compensates the inductance imparted to the resonator by the SQUID loop in the case of $\Phi_{DC}=0$ and ensures the maximum amplitude of the voltage on the resonator for an assigned [see formula (16)] amplitude of the radio-frequency signal and $\Phi_{DC}=0$:

$$u(l)=p\frac{1}{1+x_S(l)}, \tag{42}$$

where $x_S(l)$ corresponds to the values of the reactive component found by the techniques described in Sec. 3.

6. INTERPOLATION OF r_S AND x_S IN THE RANGE $0\leqslant x\leqslant 1$

Thus, we have obtained the values of the SQUID loop impedance for $x=0$ (Table II) and for $x=1$ (Table III) as functions of the fundamental SQUID parameter. We should now perform the interpolation of r_S and x_S mentioned above for all values of x in the range $0\leqslant x\leqslant 1$. Taking into account that the resonator is tuned to attain the maximum amplitude for $x=0$, we substitute (42) into (41) and also use (4). As a result, we obtain

$$r_S(x,l)+ix_S(x,l)=p(l)\left(-1-i\frac{p(l)}{1+x_S(0,l)}+\frac{1}{1-\Delta W(l)x}\exp(-i\Psi(x,l))\right)^{-1}-i. \tag{43}$$

TABLE IV.

l	$p(l)$	$\Psi(1,l)$
2	1.82	−0.13
2.5	1.37	−0.09
3	1.09	−0.03
3.5	0.92	0.02
4	0.80	0.05
4.5	0.73	0.08
5	0.68	0.10

When $x=0$, the relation (43) transforms into an identity, if $\Psi(0,l)=0$. When $x=1$, the relation (43) permits finding $p(l)$ and $\Psi(1,l)$ on the basis of the data in Table III. The calculated data are listed in Table IV.

The values obtained for $p(l)$ ensure a sawtooth law governing the variation of $W(x)$ for an assigned l, which is included as a basic experimental fact in the foundation of the present analysis. It can be concluded from the data presented that p is close to unity for characteristic values in the range $3\leqslant l\leqslant 5$. It is interesting to note that a sharp increase in the optimum value of p is observed for $l<2$. A decrease in l can be advantageous for attaining a strong output signal (Table I). However, the design difficulty in ensuring strong coupling between the loops, i.e., a value of p considerably greater than unity, imposes a constraint.

a) Impedance imparted to the resonator by the SQUID loop. The total impedance of the inductive branch of the resonator with consideration of the impedance imparted by the SQUID loop is expressed as

$$Z_L=i\omega L_T+Z_B, \tag{44}$$

where

$$Z_B(x,l)=\frac{(\omega M)^2}{\omega L_S(i+ix_S(x,l)+r_S(x,l))} \tag{45}$$

is the impedance imparted to the resonator.

We normalize Z_B with respect to the active resistance included in the inductive branch of the resonator $\omega L_T/Q$,

$$R_S(x,l)+iX_S(x,l)=\frac{Z_B(x,l)}{\omega L_T}Q. \tag{46}$$

In accordance with (44) we have

$$R_S(x,l)+iX_S(x,l)=\frac{p}{i(1+x_S(x,l))+r_S(x,l)}. \tag{47}$$

We must now make an assumption as to how the phase of the voltage on the resonator depends on the normalized constant flux. We assume that $\Psi(x)$ has the form of a power function:

$$\Psi(x,l)=\Psi(1,l)x^{\alpha}. \tag{48}$$

The only unknown in (48) is the exponent α. In addition, Φ_{DC} is an external disturbance, which causes retuning of the resonator. When the resonance processes in the oscillator circuit are considered, it can be assumed that the active component in the function of the parameter characterizing the external disturbance should have a nearly parabolic de-

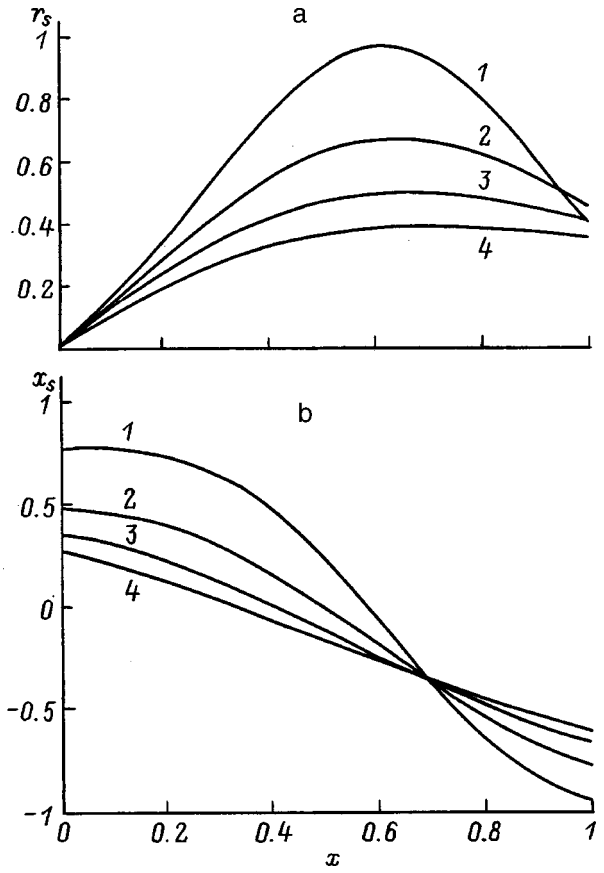

FIG. 7. Dependence of the active component r_S (a) and the reactive component X_S (b) of the impedance of the Josephson junction on the normalized bias flux for various values of the fundamental SQUID parameter $l=2$ (*1*), 3 (*2*), 4 (*3*), and 5 (*4*).

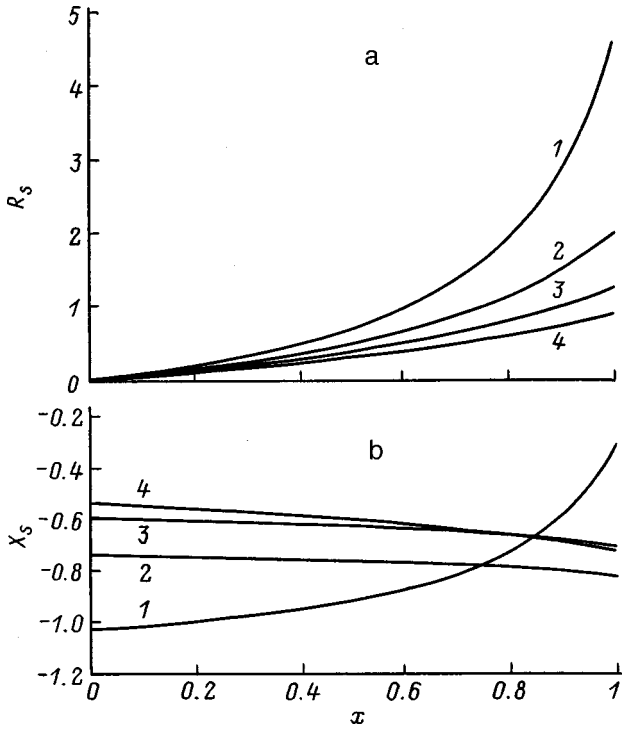

FIG. 8. Dependence of the active component R_S (a) and the reactive component X_S (b) of the imparted impedance on the normalized bias flux for various values of the fundamental SQUID parameter $l=2$ (*1*), 3 (*2*), 4 (*3*), and 5 (*4*).

pendence, while the reactive component should have a nearly linear dependence. The construction of plots of R_S and X_S with variation of α [using (47), (48), and (43)] revealed that the closest correspondence to the required course of the curves is observed for $\alpha=1$.

We can now construct the dependence of the active and reactive components on the normalized bias flux (the plots are shown in Fig. 7). Figure 8 presents plots of R_S and X_S as functions of the constant normalized bias flux for the same values of l. It can be concluded from Fig. 8 that smaller values of l correspond to larger values of R_S. This situation provides for stronger modulation of the pump amplitude on the resonator. At the same time, X_S is weakly dependent on l.

In this stage of the analysis the presence of R_N in the circuit of the Josephson junction considered within the resistively shunted junction model[1,2] can be taken into account where necessary. The resistance R_N should be connected in parallel to the junction impedance $\omega L_S(r_S+ix_S)$.

It should be recalled that a circuit containing a Josephson junction is nonlinear. The impedance as a function of Φ_{DC} found is realized at a definite pump current, which provides for the maximum value of Φ_{RF} on the boundary for the appearance of hysteresis when $\Phi_{DC}=0$. It is also important that the optimum regime for sawtooth modulation in the resonator is ensured at the value found for $p=k^2Q$, which is close to unity.

The dependence found of R_S and X_S on the constant magnetic flux passing through the SQUID loop permits the performance of a complete calculation of a microwave circuit containing a SQUID and optimization of the microwave response of the SQUID to an external disturbance in the form of a constant magnetic flux.

[1)]Here, in the equivalent circuit of a Josephson junction we have $R_N\to\infty$ (R_N is the normal resistance in the resistively shunted junction model[2]), which is valid at sufficiently low rf oscillation frequencies. Under microwave conditions R_N must be taken into account. In this case, when $x=0$, the impedance imparted to the resonator is no longer purely imaginary.

[1] M. Tinkham, *Introduction to Superconductivity* [McGraw-Hill, New York (1975); Atomizdat, Moscow (1980), 310 pp.].
[2] K. K. Likharev, *Systems with Josephson Junctions* [in Russian], MGU, Moscow (1978), 446 pp.
[3] V. N. Kornev, K. K. Likharev *et al.*, Radiotekh. Elektron. (Moscow) **26**, 2647 (1980).
[4] Y. Zhang, N. Wolters, X. H. Zeng *et al.*, *Sixth International Superconductive Electronics Conference*, Berlin (1997), Vol. 1, pp. 51–53.
[5] Y. Zhang, M. Muck, A. I. Braginski, and H. Topfer, Supercond. Sci. Technol. **7**, 269 (1994).
[6] B. Chesca, *Sixth International Superconductive Electronics Conference*, Berlin (1997), Vol. 1, pp. 54–56.
[7] T. Ryhanen, H. Seppa, R. Ilmoniemi *et al.*, J. Low Temp. Phys. **76**, 287 (1989).
[8] O. V. Lounasmaa, *Experimental Principles and Methods Below 1 K* [Academic Press, London–New York (1974); Mir, Moscow (1977), 356 pp.].
[9] V. V. Schmidt, P. Muller, and A. V. Ustinov (Eds.), *The Physics of Superconductors: Introduction to Fundamentals and Applications* [Springer, New York (1997); Nauka, Moscow (1982), 240 pp.].

Translated by P. Shelnitz

Formation of a protective film on a copper friction surface in the presence of fullerene C_{60}

B. M. Ginzburg and O. F. Kireenko

Institute of Problems in Machine Construction, Russian Academy of Sciences, 199178 St. Petersburg, Russia

M. V. Baĭdakova and V. A. Solov'ev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted June 16, 1998)
Zh. Tekh. Fiz. **69**, 113–116 (November 1999)

The influence of C_{60} additive in industrial oil on the structure of the friction surface of copper foil in a steel–copper sliding friction pair is investigated by wide-angle x-ray diffraction, scanning electron microscopy, and hardness testing. The presence of C_{60} in the lubricant leads to the formation of a thin film (of thickness <1000 Å) on the friction surface of the copper, where it protects the surface layers of the latter against major structural changes and helps to improve the tribological characteristics. © *1999 American Institute of Physics.* [S1063-7842(99)02011-5]

INTRODUCTION

Data from investigations of the tribological properties of fullerenes are very scarce and are concerned primarily with fullerene C_{60}. A few papers report studies of C_{60} in the form of a solid film used as a solid lubricant coating[1,2] and in the form of additives to fluid lubricants.[3] These few papers are enough to appreciate the promising outlook of C_{60} for the solution of various tribological problems. This conclusion can be further justified in information of a general nature on the properties of C_{60}, specifically its high elasticity and strength, low surface energy, weak intermolecular interactions, and the quasi-spherical shape of its molecules.[4–6]

In the patent literature it is reported that various modified fullerenes can be used as antifriction coatings, solid lubricants, and additives to lubricating oils.[7,8] A possible negative exception is found in fluorinated fullerenes, because some data indicate that hydrofluoric acid is formed when they are subjected to friction under atmospheric conditions.[9]

In a previous paper[10] we have demonstrated he positive influence of C_{60} used as an antiwear and antifriction additive to I-40A industrial oil on the friction characteristics of a steel–copper pair. However, the mechanisms underlying the action of C_{60} remain elusive in this case. Two hypotheses have been advanced. According to one, C_{60} molecules enter into the surface layers of the softer copper counterbody and, acting as ''molecular rollers,'' tend with comparative ease to induce rotational strain modes in the surface layers.[10] According to the second hypothesis, the high tribological characteristics of a steel–copper friction pair with a lubricant containing C_{60} additive are established by the formation of a thin protective film (presumably of tribopolymer origin) on the friction surfaces without any change in the structure of the surface layers of the material itself.[11] The objective of the present study is to ascertain by investigation of the friction surfaces the mechanisms by which C_{60} influences the processes of interfacial sliding friction of steel against copper. Since copper is an antifriction medium in the steel–copper friction pair, we have chosen copper counterbodies as the samples for our tribological tests and physical investigations.

MATERIALS

Fullerene soot was prepared by the Huffman–Krätschmer method (G. A. Dyuzhev, Project ''Duga'' of the Russian Scientific-Technical Program ''Fullerenes and Atomic Clusters,'' 1994). A mixture of fullerenes (mainly C_{60} and C_{70} in the ratio $\sim 75{:}25$) was separated from the soot by an extraction procedure using toluene as the solvent, and C_{60} (96–98% concentration) was separated from the fullerene mixture by preparative chromatography (V. P. Budtov, Project ''Khromotron'' in the same program); as a result, it was determined that the fraction of C_{60} in the fullerene soot was $\sim 10-15\%$.

The basis of the lubricant compositions was I-40A industrial oil (I-G-A-68 per State Standard GOST 17479-87), which is obtained from petroleum products and is widely used in practice for the lubrication of real friction joints.

It has been shown[12] that the addition of C_{60} to oil through a solution of C_{60} in toluene is ineffective, and the best results are obtained by adding the dry powders directly to the oil. Consequently, to prepare the lubricant compositions in the present study, we poured 5 wt. % C_{60} into the oil while mixing vigorously by mechanical means. The end result was an inhomogeneous mixture with a precipitate of coarse flocs. Before using the lubricant mixtures obtained by this procedure, we forcibly mixed them to achieve homogenization, but inhomogeneity persisted and could have been one of the factors responsible for large scatter of the experimental data.

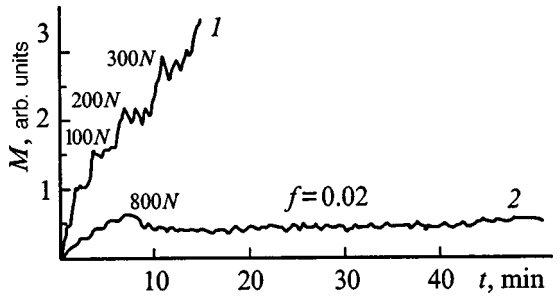

FIG. 1. Typical friction diagrams of a steel roller against copper foil with step-incremented loading of the friction joint. The friction coefficient f corresponding to the observed plateau of curve *2* is indicated.

EXPERIMENTAL PROCEDURE

As in Refs. 10–12, tribological tests were performed on a 2070 SMT-1 roller friction testing machine using a shoe-on-roller scheme. A moving roller of diameter 46 mm and width 16 mm was fabricated from wear-resistant chrome-nickel-molybdenum steel 18Kh2N4MA (State Standard 4543-71) and turned at an angular velocity $\omega = 400\ \text{min}^{-1}$, corresponding to a linear sliding velocity of 1 m/s. The shoe was wrapped tightly in copper foil (industrial grade, thickness $\sim 200\ \mu$m), which functioned as the investigated sample; the samples had a width of ~ 10 mm, and the nominal tribocontact area was $\sim 2\ \text{cm}^2$. Drops of lubricant were deposited on the surface of foil, then a load was applied to the friction joint and was incremented in steps up to pressures $\sim 4-6$ MPa, which are typical for heavily loaded friction joints.

X-ray diffraction patterns of the friction surface were recorded on a Rigaku Cor Dmax-RC single-crystal x-ray diffractometer using Ni-filtered CuK_α radiation. A scanning electron microscope was used for microscopic observations. The microhardness was measured on a PMT-3 microhardness tester.

The following types of samples were investigated: basic copper foil samples without exposure to friction; similar samples after friction treatment in pure I-40A oil; similar samples after friction treatment in oil with the addition of 5% powdered fullerene C_{60} or other materials.

EXPERIMENTAL RESULTS AND DISCUSSION

Figure 1 shows the two most characteristic types of time dependences of the friction torque (M). Curve *1* is typical of pure I-40A oil, where the increase in the friction torque, causing scoring, begins at fairly small loads (300–600 N), whereas a friction joint lubricated with oil containing C_{60} (curve *2*) can run for a long time at loads in the range 800–1200 N. In this case the friction coefficient stabilizes at a very low level, 0.015–0.030, which is unattainable using the pure oil or the same oil with other carbon additives.

In addition, when C_{60} is present in the oil, almost the entire friction surface of the copper acquires a mirror sheen, which is indicative of the formation of a wear-resistant structure on it,[10] there are no traces of copper transferring to the steel roller, and debris is not observed in the lubricant or on the friction surface. Subsequent wiping of the roller or replacement of the lubricant by the pure basic oil does not disrupt the stable operation of the friction joint, and the friction coefficient remains at the same low level as in the presence of C_{60}. This result indicates the formation of an anti-friction, wear-resistant structure on the friction surface.

Table I gives the relative intensities of x-ray diffraction from various crystallographic planes for a number of typical samples of the investigated copper foil. For the basic samples the ratios of the intensities of different reflections (the intensity of the strongest reflection was set equal to 100 arb. units) differ appreciably from the analogous ratios for a standard sample of polycrystalline copper with randomly oriented grains.[13] This means that the surface of the foil has a certain crystalline, grain-oriented texture.

For samples subjected to friction with pure I-40A oil (even in the ''wear-resistant regime'' created by very careful and gradual escalation of the normal load and by patient wearing-in with each new load), an abrupt redistribution of the intensities of the principal diffraction peaks is observed, tending to the pattern of disoriented polycrystalline copper, which implies breakdown of the texture. This behavior is consistent with the results of an investigation of the low-temperature transverse fractures of such samples by scanning electron microscopy: Beneath the friction surface is observed

TABLE I. Intensities of principal diffraction peaks from various crystallographic planes of copper in relative units.

	Reflection indices and their angular positions					
	111	200	220	311	400	331
Sample	43.3°	50.42°	74.08°	89.87°	116.84°	136.54°
Standard	100	46	20	17	3	9
B	1	100	5	<1	<1	<1
O	90	100	100	<1	<1	<1
F	5	100	30	6	<1	4
F	1	100	17	8	2	6
F	9	100	11	9	3	8
F	17	100	8	3	<1	3
S + F	4	100	3	<1	<1	<1

Note: B) Basic sample; O) sample working in I-40A oil in the wearless friction regime; F) sample working in the same oil with 5% C_{60}; S+F) sample working in the same oil with 5% fullerene soot. The F samples differ in the load on the friction joint and the working time in the joint. Recording of the diffraction patterns along and across the rolling direction of the foil gives essentially identical results.

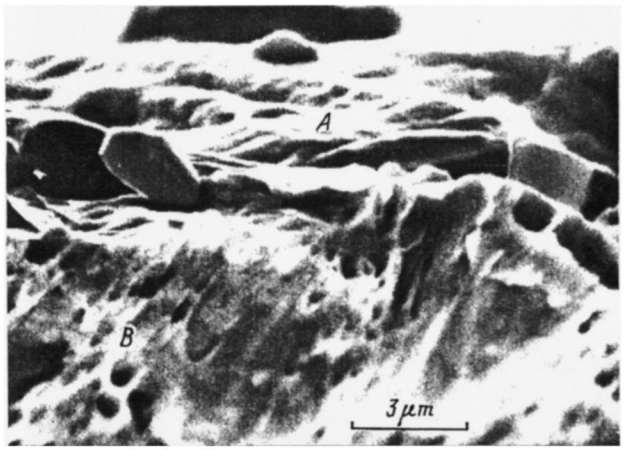

FIG. 2. Electron micrograph of low-temperature fracture of copper foil after friction with I-40A industrial oil with a 5% Co_{60} additive: *A* — section of the friction surface coated with a thin (not more than 1000 Å) film formed during the friction process; *B* — fracture zone with unaltered structure.

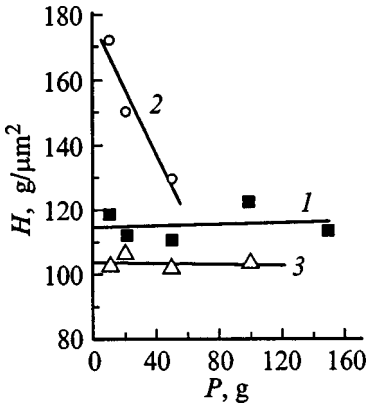

FIG. 3. Dependence of the microhardness *H* of the surface of copper foil on the indenter load *P*: *1* — original surface of the foil before friction; *2* — surface of the foil after friction with lubricating oil containing C_{60}; *3* — surface of the foil after friction with lubricating oil in the so-called wear-resistant regime.

a layer having a finely disperse structure; this layer is usually a few micrometers thick. Consequently, x-ray diffraction and scanning electron microcopy disclose major changes in the structure of the copper surface layers as a result of friction with lubricating oil when paired with a steel counterbody, even in the so-called wear-resistant friction regime.

A fundamentally different pattern is observed for samples tested with a C_{60}-containing lubricant. Table I gives typical x-ray diffraction data for several samples subjected to friction under the stated conditions for various durations and at various loads. For all the investigated samples only comparatively small changes in the diffraction patterns from those obtained for the basic samples are observed (the ratios of the intensities of different reflections and their positions and widths). Consequently, the presence of C_{60} in the lubricant prevents friction, in effect, from disrupting the original surface structure of the copper foil, even after a long testing period (up to several hours) and large loads (up to 1000 N).

The electron microscope examinations show that, owing to friction in the presence of C_{60}, the copper friction surface retains its constant basic structure and is coated with a film of thickness no greater than 1000 Å (Fig. 2). Here the friction surfaces even retained a system of striations (oriented in the rolling direction), which vanished completely with the use of ordinary lubricating oil in the so-called wear-resistant regime. Thus, the energy dissipation mechanism associated with friction in the presence of C_{60} is not related to structural modification of the material itself in the surface layers, but in all probability depends only on the properties of the indicated extremely thin film formed on the surface of the copper.

The microhardness of the friction surface of samples subjected to friction in the presence of C_{60} is higher than the microhardness of the original surface (cf. curves *1* and *2* in Fig. 3). In ''wearless friction'' with lubricating oil the formation of a finely disperse and probably more friable surface layer tends to lower the microhardness (curve *3* in Fig. 3). Also, curves *1* and *2* disclose an insensitivity of the microhardness to variation of the indenter load. This fact implies homogeneity of the mechanical properties in the direction perpendicular to the surface within the indenter penetration limits. On the other hand, the microhardness of the friction surface formed in the presence of C_{60} decreases as the load is increased, indicating the presence of a thin, but harder surface layer than the base material. Two structural modifications of the surface layers can be postulated on the basis of these data: either a single surface film or a strengthened copper layer (as a result of cold hardening, the infiltration of harder particles, from the steel counterbody for example, into the surface, etc.). However, an x-ray structural analysis of the surface shows that the second structural alternative does not occur; when oil with a C_{60} additive is used, x-ray examination of the friction surfaces does not reveal any shifting or broadening of the principal reflections, and the texture changes very little or not at all.

CONCLUSION

The sum total of the results obtained by three methods (scanning electron microscopy, wide-angle x-ray diffraction, and microhardness testing) shows that for friction in the presence of C_{60} a thin ($\sim$1000 Å) film is formed on the surface of a copper counterbody, protecting the friction surface against wear and also lowering and stabilizing the friction coefficient.

An analysis of the literature suggests several hypotheses to account for the composition and structure of this film. We first carry out an elementary calculation to show whether the concentration of C_{60} is sufficient for it to form a continuous layer on the friction surface layer. The area of the friction surfaces of the investigated samples was $\sim 2\ cm^2$, and the actual thickness of the gap between them was $\sim 1\ \mu m$. For a $\sim$5% concentration of C_{60} in the oil the total number of molecules in the gap is $\sim 0.8\times 10^{16}$. If the C_{60} molecules are hexagonal close-packed, the area associated with one molecule is $\sim 100\ Å^2$. Accordingly, all the molecules in the gap can cover the area of the friction surface with $\sim$40 monolayers with close packing in the layer. The estimated 40-monolayer thickness ($\sim$400 Å) is reasonably consistent with

the electron microscope data ($\leqslant 1000$ Å). The concentration of C_{60} is therefore sufficient for it to form a monolithic film. However, when the foil is kept for several hours in toluene, which is the best solvent for C_{60}, this film is not removed from the surface. Consequently, it certainly could not consist of molecules of pure unmodified C_{60}.

Several other possibilities exist. We know that under high pressure, even without added shear stresses, C_{60} can form polymer films,[14] structures of the diamond type,[15] or amorphous carbon.[16] Shear strains in the tribocontact zone can be conducive to the onset of these processes at even lower temperatures and pressures than in the cases described here. On the other hand, the presence of C_{60} in oil can result in the formation of a tribopolymer film. This hypothesis is supported by the results of an investigation of the tribological properties of C_{60} additives with grafted polystyrene chains; they have yielded even better results than pure fullerene C_{60} (Ref. 17).

We also note that fullerene soot (before the extraction of fullerenes from it), used as an additive to I-40A oil, yields essentially the same results as C_{60} in tribological tests, structural examinations (see Table I), and hardness tests. With the addition of soot to the oil (5%) the C_{60} content in it is 0.5–0.75%. The positive effect of such small quantities of C_{60} suggests that C_{60} can act as a catalyst for the polymerization of hydrocarbons present in petroleum oils. Films having a diamond or amorphous carbon structure or the inclusion of particles of these solid phases in the tribological surface film can then be attributed to the presence of a layer that is thin but harder than the base material on the friction surfaces of copper.

[1] B. Bhushan, B. K. Gupta, Van G. W. Cleef, C. Capp, and J. V. Coe, Tribol. Trans. **36**, 573 (1993).
[2] B. Bhushan, B. K. Gupta, Van G. W. Cleef *et al.*, Appl. Phys. Lett. **62**, 3253 (1993).
[3] B. K. Gupta, B. Bhushan, Lubr. Eng. **50**, 524 (1994).
[4] A. V. Eletskiĭ and B. M. Smirnov, Usp. Fiz. Nauk **163**(2), 33 (1993) [Phys. Usp. **36**, 202 (1993)].
[5] R. F. Curl and R. E. Smalley, Sci. Am., No. 265, pp. 54–63 (1991).
[6] M. Nuñez-Regueiro, J. L. Hodeau, L. Abello, and C. Lucazeau, Diamond Relat. Mater. **2**, 1203 (1993).
[7] M. Taniguchi, Y. Tomioka, N. Kumegava, and M. Ishibashi, Jpn. Patent No. 05-179269 (1991).
[8] A. O. Patil, G. W. Schriver, and R. D. Lundberg, U.S. Patent No. 5292444 (1992).
[9] R. Taylor, A. G. Avent, T. J. Dennis *et al.*, Nature (London) **355**, 27 (1992).
[10] B. M. Ginzburg, O. F. Kireenko, D. G. Tochil'nikov, and V. P. Bulamov, Pis'ma Zh. Tekh. Fiz. **21**(23), 35 (1995) [Tech. Phys. Lett. **21**, 966 (1995)].
[11] *Report on Project ''Tribol'' of the Russian Scientific-Technical Program ''Fullerenes and Atomic Clusters''* [in Russian] (St. Petersburg, 1996).
[12] B. M. Ginzburg, D. G. Tochil'nikov, O. F. Kireenko, and V. P. Bulamov, Pis'ma Zh. Tekh. Fiz. **21**(22), 62 (1995) [Tech. Phys. Lett. **21**, 933 (1995)].
[13] Joint Committee on Powder Diffraction Standards (JCPDS), International Centre for Diffraction Data, Swarthmore (USA), No. 4-836 (1993).
[14] I. O. Bashkin, V. I. Rashcupkin, A. F. Gurov *et al.*, J. Phys.: Condens. Matter **97**, 7491 (1994).
[15] O. G. Epanchintsev, A. S. Zubchenko, Yu. D. Tret'yakov *et al.*, Dokl. Akad. Nauk **330**, 201 (1993).
[16] H. Hirai, K. Kondo, N. Yoshizawa, and M. Shiaishi, Appl. Phys. Lett. **64**, 1797 (1994).
[17] B. M. Ginzburg, V. N. Zgonnik, D. G. Tochil'nikov *et al.*, Pis'ma Zh. Tekh. Fiz. **22**(14), 1 (1996) [Tech. Phys. Lett. **22**, 554 (1996)].

Translated by James S. Wood

Interaction of C_{60} molecules with a (100) Mo surface

N. R. Gall', E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted August 19, 1998)
Zh. Tekh. Fiz. **69**, 117–122 (November 1999)

The adsorption, initial stages of film growth, and transformation of an adlayer of C_{60} molecules on a (100) Mo surface upon heating are studied under ultrahigh-vacuum conditions. It is shown that the C_{60} molecules remain intact in the adsorbed state up to 760 K. Layer-by-layer growth of a fullerite film is observed at room temperature, while tower-shaped crystallites grow up from a loosely packed monolayer with an approximate concentration of C_{60} molecules equal to $(1.3\pm 0.2)\times 10^{14}$ molecules$\cdot$cm^{-2} at 500–600 K. In the latter case the percentage of the surface occupied by them depends on the temperature and the impinging flux density of fullerene molecules, but after a certain stationary value has been achieved, it scarcely depends on the exposure time. [S1063-7842(99)02111-X]

INTRODUCTION

Fullerenes represent a new, relatively recently discovered form for the existence of carbon. Unlike such widely known allotropic forms as diamond, which is a three-dimensional crystal of the cubic system, graphite, which consists of a set of two-dimensional planes weakly bound to one another, and carbyne, which consists of long carbon filaments, fullerenes are generally individual molecules of strictly definite size and composition, such as C_{60}, C_{70}, C_{84}, C_{100}, and other molecules of this homologous series. All fullerene molecules consist only of carbon atoms, which form a closed cage enclosing a cavity free of matter.

Among the fullerenes, the C_{60} molecule is most accessible to study. The C_{60} molecule has the form of a sphere with a diameter drawn through the centers of carbon atoms comprising it equal to 7.08 Å.

For practical and scientific purposes we must understand how fullerene molecules interact with matter, particularly with solid surfaces. The basic study of the adsorption of C_{60} molecules on the surfaces of solids is a fairly complex problem, mainly because of the complex structure of the object of study. Nevertheless, there have been at least several dozen reports of theoretical and experimental studies performed by various methods. Silicon,[1–5] noble metals,[6–8] and refractory metals[9] are usually employed as substrates.

We previously studied the adsorption of C_{60} on (111) Ir (Refs. 10 and 11) and $(10\bar{1}0)$ Re surfaces (Refs. 12 and 13). The contact and thermal stability of C_{60} molecules on surfaces of these metals were determined, and the transformation of the adlayer of C_{60} molecules upon substrate heating were determined. It was found that the adsorbed C_{60} molecules maintain their fullerene nature up to a certain temperature, which depends on the type of substrate ($\sim$800 K on Re and $\sim$1000 K on Ir). At higher temperatures the molecules decompose, and the carbon released in the adlayer dissolves in the bulk of the substrate (on Re) or graphitizes (on Ir). A very interesting effect known as intercalation, i.e., the spontaneous penetration of C_{60} molecules adsorbed on a monolayer graphite film to the iridium[11] beneath the film, was observed for the first time. A practical device, viz., an ultrahigh-vacuum regenerator for creating a flow of carbon atoms free of carbon clusters, was created on the basis of our experiments.

In Refs. 10 and 11 we discovered that the C *KVV* Auger peak for C_{60} molecules has an energy of 269 eV, which is almost 3 eV lower than the value, say, in graphite, metal carbides, and adsorbed carbon clusters. This Auger peak energy is observed both for thick films of adsorbed C_{60} molecules (5–10 monolayers) and for submonolayer coatings. The utilization of this phenomenon for distinguishing fullerenes adsorbed on a surface from other possible states of carbon was proposed.

In the present work we studied the adsorption of C_{60} molecules on a molybdenum surface, which is widely used in electronics, in vacuum technology, and in the technology for growing diamond films. We previously conducted a detailed study of the processes occurring when carbon interacts with this surface and had fairly exact knowledge of the temperature ranges for the beginning of its dissolution, the conditions for the formation of surface and bulk carbides, etc.[14,15] In addition, molybdenum has intense Auger peaks with energies of 186 and 221 eV, which renders it extremely convenient for studying the adsorption of strongly screening materials, such as C_{60} molecules.

SAMPLES AND EXPERIMENTAL METHOD

The experiments were carried out in the high-resolution Auger spectrometer described in Ref. 16 under ultrahigh-vacuum ($P\sim 10^{-10}$ Torr) conditions. We were able to record the Auger peaks directly from heated samples in the range 300–2100 K. The samples used were annealed molybdenum ribbons measuring $0.02\times 1\times 40$ mm, which were cleaned and characterized according to a standard procedure. The ribbons were textured, and a (100) face emerged on the surface. The surface work function was $e\varphi = 4.40$ eV, and the

surface was homogeneous with respect to $e\varphi$. According to the data from electron-microscopic investigations, the mean grain diameter was $\sim 15\mu$m.

The fullerenes were spray-deposited onto the entire surface of the ribbon at a 65° angle to a normal from a Knudsen cell. The cell was charged with a weighed portion of C_{60} molecules (~ 0.05 g) of 99.5% purity. After activation, the cell provided a stable and easily regulated flow of fullerene molecules with a flux density $\nu_C = 10^{10} - 10^{13}$ molecules $\cdot \mathrm{cm}^{-2} \cdot \mathrm{s}^{-1}$.

The study of the adsorption of fullerenes on metal surfaces is greatly complicated by the fact that such a surface, even if initially clean, is unavoidably contaminated by carbon released in the adlayer upon decomposition of the molecules during experiments. In our experiments the surface was cleaned by heating to $T = 2100$ K after each series of measurements. As a result, all the undesorbed C_{60} molecules decomposed into atoms, the carbon dissolved in the bulk of the substrate, and only a surface carbide (MoC) was present on the surface in a concentration equal to 1×10^{15} atoms $\cdot \mathrm{cm}^{-2}$ (Refs. 14 and 15). The surface carbide contains strongly bound carbon atoms, which are removed from the Mo surface only at $T > 1400$ K by dissolving them in the bulk of the metal. When the temperature is lowered to $T < 1400$ K, carbon atoms escape from the bulk emerge to the surface and reform the surface carbide. It is logical to assume that their concentration remains unchanged when fullerene molecules are adsorbed and that the Auger signal emitted by these atoms is screened to the same degree as the Auger signal of molybdenum. This assumption is justified by the proximity of the two Auger signals ($E_{Mo} = 221$ eV, $E_C = 273$ eV). In order to reveal the ''pure'' Auger signal from the fullerene molecules, the contribution originating from the surface carbide was subtracted. The value used for this purpose was the initial amplitude (before the adsorption of C_{60} molecules), which was reduced to the same degree that the substrate Auger signal decreased. The question of whether this contribution is subtracted or not subtracted is stipulated in each specific case.

ABSOLUTE CALIBRATION OF THE FLUX OF C_{60} MOLECULES

The intensity of the Auger signal of carbon from C_{60} molecules by itself does not provide any information on the number of these molecules present on the surface, since the C atoms in an adsorbed molecule screen one another in a complicated manner. To calibrate the number of molecules, we determined the degree of screening of the substrate Auger signal corresponding to a particular number of C_{60} on the surface. For this purpose, C_{60} molecules were deposited on the surface of an auxiliary iridium ribbon, which was placed in the Auger spectrometer parallel to the working molybdenum ribbon, in a small, presumably submonolayer (according to the data in Ref. 10) concentration. They lowered the amplitude of the Ir Auger peak at $E = 169$ eV by a factor of 1.6. Then the ribbon was heated to $T = 1700$ K. As we previously showed, the fullerene molecules then decompose, the carbon released actively migrates over the surface, is distributed equally on the two sides of the ribbon, and graphitizes. We measured the Auger signal of the graphite formed, and it amounted to less than 1/2 of a monolayer. Since the concentration of carbon in a graphite monolayer $N_C = 3.9 \times 10^{15}$ cm^{-2}, there was initially $\sim 3.9 \times 10^{15}$ C atoms or 6.5×10^{13} C_{60} molecules per cm^2 on the surface. Assuming that the same number of fullerene molecules screens the molybdenum and iridium Auger signals by the same factor (the energies of both Auger signals, which are equal to 169 and 221 eV, are very close), we can calibrate submonolayer coverages of C_{60} molecules on molybdenum. Assuming that the accommodation coefficient of the molecules on the surface is close to unity (this assumption is supported by the large values of the sublimation energies of C_{60} molecules adsorbed on various substrates), we can calibrate the flux impinging on a surface. In the present example it was $\sim 6.5 \times 10^{11}$ molecules $\cdot \mathrm{cm}^{-2} \cdot \mathrm{s}^{-1}$.

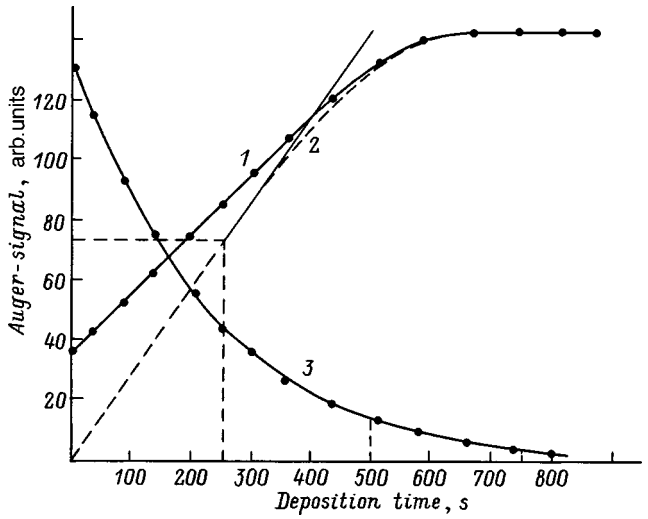

FIG. 1. Deposition of C_{60} on (100) Mo at room temperature with a flux density $\nu_{C_{60}} = 6.5 \times 10^{11}$ molecules $\cdot \mathrm{cm}^{-2} \cdot \mathrm{s}^{-1}$: *1* — carbon Auger signal, *2* — carbon Auger signal after subtraction of the contribution of the surface carbide, *3* — molybdenum Auger signal.

ADSORPTION AT ROOM TEMPERATURE

Figure 1 presents the variation of the carbon and molybdenum Auger signals during the adsorption of a constant flux of C_{60} atoms on (100) Mo at room temperature. Curve *2* presents the carbon Auger signal, from which the contribution of the surface carbide has been subtracted. As we see, the molybdenum Auger signal decreases monotonically and drops to zero, while the carbon Auger signal increases almost linearly up to a deposition time of ~ 250 s. Then the rate of its increase slows, and it reaches saturation very abruptly. This means that after $t = 250$ s, the newly arriving C_{60} molecules screen not only the substrate, but also the carbon Auger signal from previously adsorbed molecules. Thus, it can be stated that filling of the first monolayer was completed and the growth of the next monolayers began at $t = 250$ s.

What is a monolayer of C_{60} molecules on (100) Mo? We note that the first monolayer screens the substrate Auger signal quite strongly, i.e, by a factor of ~ 2.9! Similar screening

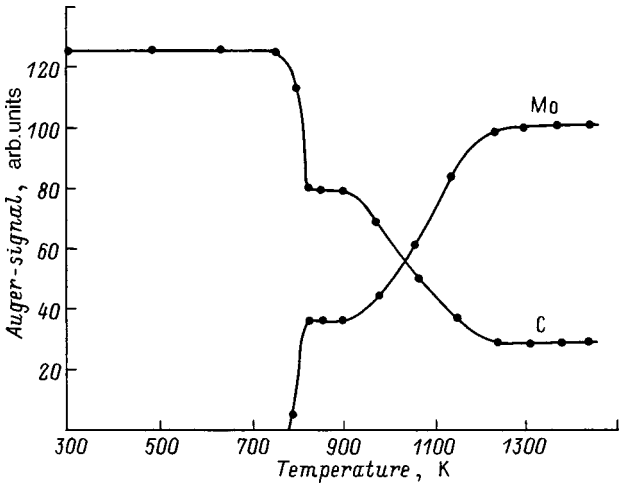

FIG. 2. Annealing of a thick film of C_{60} molecules on (100) Mo. Initial state — film of C_{60} molecules deposited at room temperature. The thickness of the film is ~6.5 layers.

is provided by a film containing ~2.5 graphite monolayers, and the total number of carbon atoms in such a film is 1.0×10^{16}cm^{-2}. Is this possible? If we assume for a rough estimate that our adlayer contains approximately the same number of carbon atoms, which are, however, joined in C_{60} molecules, there should be $\sim 1.6 \times 10^{14}$ molecules per cm^2 on the surface.

On the basis of the absolute calibration of the flux that we made, we can state that $\sim 1.63 \times 10^{14}$molecules$\cdot$cm^{-2} have impinged on the surface after 250 s, in good agreement with the estimate presented. Let us try to understand how such a number of molecules can be accommodated on a surface in one layer. The diameter of the C_{60} molecule drawn through the centers of its atoms is 7.1 Å, and the structural-chemical diameter that we need is greater by the covalent diameter of a C atom or by 1.4 Å, i.e., is equal to 8.5 Å.[17] The concentration of close-packed spheres with such a diameter on a surface is $N_{\text{cl-packed}} = 1.64 \times 10^{14}$molecules$\cdot$cm^{-2}, which closely coincides with the estimate given above and allows us to assume that a close-packed monolayer of C_{60} molecules grows on (100) Mo.

How does further growth of the film take place? Some estimates can be made from the Auger-spectroscopic data. We assume that the concentration of C_{60} molecules in the second layer is the same as in the first, i.e., $\sim 1.64 \times 10^{14}$molecules$\cdot$cm^2. Then, another 250 s are needed to deposit this layer, and the screening of the substrate after completion of its construction should be by a factor of $3 \times 3 = 9$. As is seen from Fig. 1, at a deposition time of 500 s, the substrate is screened by a factor of 9.4, and at a deposition time of 750 s (which corresponds to the formation of a third layer) it is screened by a factor of 26.5. These data attest to the layer-by-layer growth mechanism of a fullerite film on (100) Mo at room temperature.

ANNEALING OF A FULLERITE FILM ON (100) Mo

Figure 2 presents the variation of the carbon and molybdenum Auger signals during the annealing of a thick fullerite film. The estimated film thickness is ~6.5 monolayers. It is seen that up to 750 K the Auger signals of the adsorbate and substrate do not vary. The energy of the carbon Auger peak is 269 eV, providing evidence that specifically C_{60} molecules are present on the surface. In the range 750–820 K the film rapidly decomposes: the carbon Auger signal weakens, while the substrate Auger signal intensifies. At the temperatures indicated even individual carbon atoms do not dissolve in the bulk of a molybdenum substrate.[15] The C_{60} molecules are deposited uniformly over the entire ribbon surface: the escape of particles as a result of migration is unlikely. Therefore, the only plausible reason for the departure of C_{60} is thermodesorption.

Up to 900 K the surface is still covered by some carbon, which screens the substrate Auger signal by a factor of 2.7. This carbon is not in fullerene molecules, since the energy of the Auger peak is 272 eV. The bulk of the C_{60} molecules have probably desorbed, and only a layer of residues of the carbon cages of fullerene molecules remains on the surface. At $T > 900$ K the carbon from this layer begins to dissolve in the bulk of the molybdenum, and at $T > 1300$ K there is only a surface carbide on the surface. As we previously showed in Refs. 14 and 15, at these temperatures the form in which carbon was adsorbed on the surface is already unimportant, since all the molecules and their residues have decomposed and there are only individual C atoms in the adlayer.

It was important to understand up to what temperature the individual adsorbed molecules of C_{60} remain intact. For this purpose, a small number of C_{60} molecules was deposited on the surface, and the adlayer obtained was annealed. It was found that the individual C_{60} molecules do not decompose upon adsorption on Mo up to $T \leqslant 760$ K and maintain their structure at lower temperatures.

DEPOSITION AT ELEVATED TEMPERATURES

The preceding experiments allowed us to single out the temperature range in which we could refer to the adsorption of fullerene molecules proper, rather than fragments of them, on MO. The cutoff temperature for the experiments was 760 K, at which C_{60} molecules are known not to decompose on the surface.

As we have already seen, when the thickness of the adlayer is greater than one monolayer, the degree of screening of the substrate Auger signal becomes more informative than the intensity of the adsorbate Auger signal, which varies slightly as the second and ensuing layers grow. Figure 3 presents the variation of the molybdenum Auger signal as C_{60} molecules are deposited at various temperatures by the same flux. As is seen, all the curves coincide in the initial stages, but then they diverge. The curves recorded at each temperature each reach their own characteristic plateau, and then vary very weakly. It is noteworthy that the real differences between the thicknesses of the adlayers represented by, say, curves *3* and *4* are quite considerable and can fully be of the order of half of a monolayer or more. At the same time, the differences in deposition temperature amount to only 40 K!

The physical picture of the processes was absolutely un-

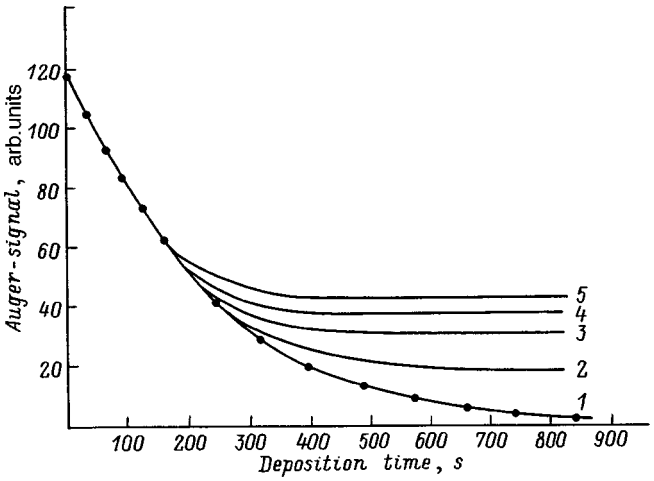

FIG. 3. Variation of the molybdenum Auger signal during the deposition of C_{60} molecules on (100) Mo at various temperatures T, K: *1* — 300, *2* —600, *3* — 645, *4* — 680, *5* — 700. The flux density is the same as in Fig. 1.

intelligible, and in order to elucidate it, we altered the experiment: the Auger spectra were recorded in a flow of fullerene molecules impinging on the surface at once at high temperatures. Figure 4 shows the corresponding stationary states of the carbon and substrate Auger signals corresponding to the plateaus in Fig. 3. In this experiment C_{60} molecules were deposited at $T=760$ K until a stationary state was achieved. Then, the temperature was lowered, the next stationary state was recorded, and so forth and so on.

It is seen that there is a plateau in the range 760–700 K (curve *1* in Fig. 4), where the number of C_{60} molecules accumulated on the surface does not depend on temperature. The concentration of molecules in this state is such that the substrate Auger signal is screened by a factor 2.5. On the basis of the data presented in Fig. 1, we estimated this concentration to be $\sim(1.3\pm0.2)\times10^{14}$molecules$\cdot$cm^{-2}, i.e., close to the value for coherent close-packed covering of an unreconstructed (100) Mo face. Upon heating above 800 K, C_{60} molecules are not desorbed from this state, but they decompose, and their fragments remain in the adlayer. At the same, at temperatures below 760 K, according to the energy

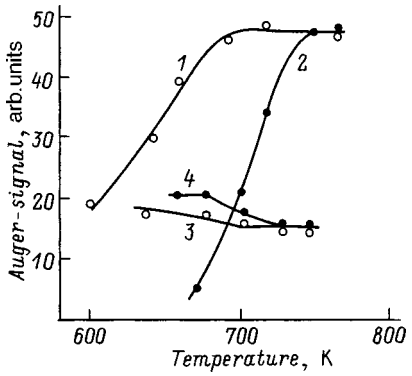

FIG. 4. Stationary values of the carbon and molybdenum Auger signals during the deposition of C_{60} molecules at elevated temperatures: *1,2* — molybdenum Auger signal; *3,4* — carbon Auger signal (reduced 2×); $\nu_{C_{60}}$, molecules$\cdot$cm$^{-2}\cdot$s^{-1}: *1,3* — 6.5×10^{11}, *2,4* — 6.9×10^{12}.

of the Auger peak, the fullerene molecules remain intact on the surface. We called this state a high-temperature monolayer. If the temperature is now lowered, say, by 20 degrees, a different stationary state characterized by its own degree of screening and its own intensity of the carbon Auger signal is achieved very quickly. We lower the temperature again by a small amount, a new state is achieved, and so forth and so on. If the flux of fullerene molecules is increased, the same stationary states are achieved at higher temperatures (Fig. 4, curve *2*).

What is the nature of these states? Our first thought is that this may be an equilibrium curve, and everything is determined simply by an adsorption-desorption equilibrium on the surface. Then let us switch off the flow or simply raise the temperature, and we should be able to move back along the curve. However, nothing happens in either case; switching off the flow or raising the temperature (in the range up to 760 K, of course) does not alter the states achieved in any way, and they are, as it were, frozen. If the temperature is raised above 800 K, the excess C_{60} molecules are desorbed, and only a standard layer of their fragments remains on the surface. Thus, the curve shown in Fig. 4 has a nonequilibrium character.

What does this mean? When we achieve our stationary states in a flow of molecules, either the newly arriving molecules are desorbed from the surface or we simply do not detect them. The following experiment was set up to resolve this dilemma. The desorption of C_{60} molecules from one such state, which was attained at 645 K, was studied under conditions where deposition was continued for 4 min in one case and for 25 min in the other case. It was found that the desorption time of molecules from the surface in the two cases just described differ significantly: while in the former case achievement of a high-temperature monolayer of C_{60} molecules occurs after 1.5–2 min, in the latter case a period several times longer, i.e., ~20 min, is required! Such prolongation of the desorption process is evidence that C_{60} molecules accumulate in the adlayer, but Auger spectroscopy does not detect their presence.

FORMATION OF THE ADLAYER DURING HIGH-TEMPERATURE DEPOSITION

In our opinion, all the laws observed can be explained, if we assume that crystallites of C_{60} molecules grow on the surface. The model described is shown schematically in Fig. 5. The crystallites grow up from the high-temperature monolayer in the form of towers. Each of the experimentally detected stationary states corresponds to its own percentage of the area which is occupied by crystallites. Exposure to a constant flux scarcely leads to any increase in the cumulative area, but does lead to growth of the towers in height. However, lowering the temperature or increasing the flux at once leads to the nucleation and growth of new towers and, consequently, in a decrease in the area free of them.

Let us consider what the Auger signals of carbon and the substrate will be in the case of the realization of such a mechanism. The carbon Auger signal will be determined by the emission from the 2–3 uppermost layers of each tower

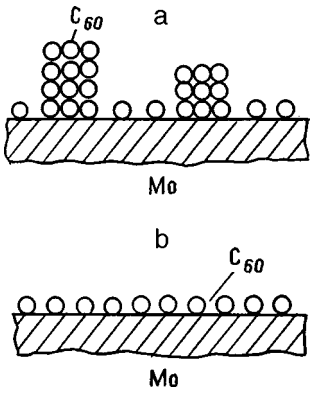

FIG. 5. Hypothetical scheme of the structure of adlayers formed during the high-temperature deposition of C_{60} molecules. a: Tower-shaped crystallites; b: high-temperature monolayer.

and will not depend on their height. The substrate Auger signal will be observed only from regions free of towers and will be strictly proportional to the area of those regions.

However, Auger electron spectroscopy is only an indirect method for determining the surface topography. Although we could not suggest any other models which would describe all the experimental observations, more direct measurements were still needed to corroborate it. We decided to perform them using atomic force microscopy (AFM). Tunneling microscopy is not suitable, since a fullerite film is an insulator.

A sample was prepared, on which C_{60} molecules were adsorbed at $T=620$ K for 20 min from a stream with a flux density $\nu_{C_{60}}=1\times10^{13}$ molecules$\cdot$cm$^{-2}\cdot$s^{-1}. The measurements were performed under room conditions on an instrument which provides nearly atomic spatial resolution.

The results of the AFM measurements are presented in Fig. 6. The crystallites formed in the shape of towers with flat tops and deep gaps between them are clearly visible. The dimensions of the crystallites are similar in both area and height, their mean transverse dimension is ~3000 Å, and their concentration on the surface is $\sim10^9$cm^{-2}. It is easy to see that the crystallites occupy $\sim80\%$ of the area of the sample. It noteworthy that, by its nature, AFM cannot detect deep narrow gaps, if their depth exceeds their width, since the cantilever serving as a probe is sharpened to a 28° angle and gives a picture of the type shown in Fig. 6 when deep and narrow gaps are scanned.

It seemed tempting to calculate the area occupied by crystallites for various values of the temperature and the flux density of C_{60} molecules impinging on the surface from the Auger-spectroscopic data. This area can be determined on the basis of the substrate Auger signal, since Mo Auger electrons can escape unabsorbed only from an area free of crystallites. The calculations were performed using the formula

$$S_{cr}=1-I_{Mo}/I^0_{Mo},$$

where S_{cr} is the fraction of the area occupied by crystallites, and I_{Mo} and I^0_{Mo} are the current and initial (beneath the high-temperature monolayer) values of the molybdenum Auger signal.

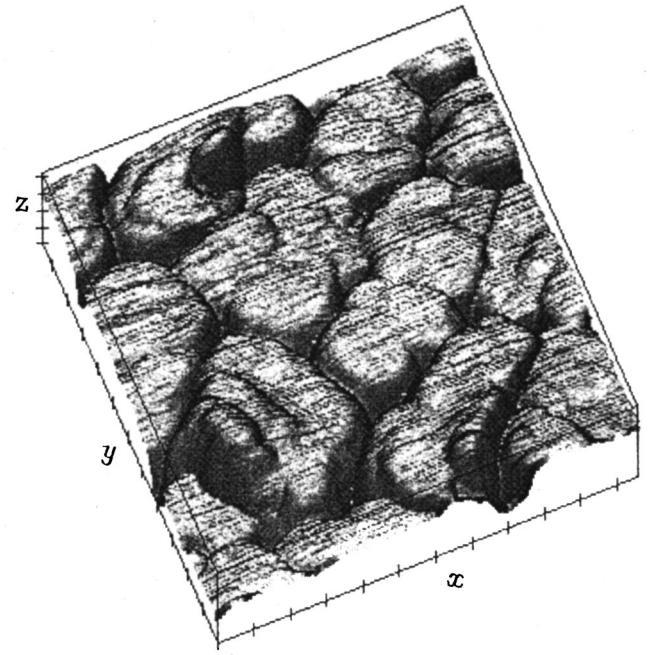

FIG. 6. Microscopic image of a portion of a (100) Mo surface with crystallites grown on it. One scale division along the *x* and *y* axis — 100 nm, one scale division along the *z* axis — 10 nm.

According to the results of the calculations, the fraction of the area occupied by crystallites on the sample used for the AFM measurements was $\sim85\%$. As we see, this value is close to the percentage actually observed on the microscopic image of the surface (compare with Fig. 6).

DISCUSSION OF RESULTS

It would be interesting to compare the results obtained on molybdenum with the results obtained on other substrates. We have studied the absorption of C_{60} molecules on Ir, Re, and Mo, as well as on a graphite monolayer grown on Ir, Re, and Mo, and obtained preliminary data on W, Si, and Ni surfaces. The literature contains a great deal of data regarding adsorption on silicon and noble metals. We note that the general character of the laws noted above is preserved on all metal substrates. However, there are significant differences. For example, C_{60} molecules remain intact on Ir, Re, Mo, Si, and Ni at room and even higher temperatures, while they decompose on tungsten already at 300 K. Crystallites form already at room temperature on silicon and nickel and possibly on rhenium, while layer-by-layer film growth probably occurs on Mo and tungsten. On Mo, Ir, and Re the monolayer is probably a nearly close-packed layer of molecules and does not depend on the crystal geometry of the substrate. There is presently no physical theory which would permit predicting the type of adsorption of fullerene molecules on a particular substrate *a priori*, and experimental study remains the most important source of information on these processes.

BRIEF RESULTS AND CONCLUSIONS

It can be said that a very complete physical picture of the processes taking place upon the adsorption of C_{60} molecules

on a Mo surface at temperatures from 300 to 2000 K or more has emerged from the experimental data obtained. Molybdenum is the only substrate on which it has been possible to analyze the picture of the processes taking place in such great detail.

The adsorption of C_{60} molecules at room temperature leads to the layer-by-layer growth of a fullerite film. The film obtained is thermostable up to $T=760\,\mathrm{K}$, above which most of the fullerene molecules are desorbed, while the remaining molecules with a concentration of the order of $1.3\times10^{14}\mathrm{molecules}\cdot\mathrm{cm}^{-2}$ decompose. Their fragments are found on the surface up to $T=900\,\mathrm{K}$, above which carbon begins to dissolve in the molybdenum bulk. Individual molecules are adsorbed on a clean surface without decomposing up to 760 K, but at higher temperatures adsorption leads to their decomposition.

The deposition of C_{60} molecules at elevated temperatures in the thermostability region ($<760\,\mathrm{K}$) leads to the growth of a structurally unusual film in the form of tower-shaped crystallites. The number of crystallites depends on temperature and the flux density of the impinging molecules, but their total area after the achievement of a certain stationary value depends on the deposition time of C_{60} molecules.

We thank A. N. Titkov, I. V. Makarenko, and Z. Vakar for performing the AFM measurements.

This work was carried out with the support of the Russian State Program ''Fullerenes and Atomic Clusters,'' Project No. 98060.

[1] Xu Hang, D. M. Chen, and W. N. Creager, Phys. Rev. Lett. **70**, 1948 (1993).
[2] Y. Z. Li, M. Chander, J. C. Partin, and J. H. Weaver, Phys. Rev. B **45**, 13 837 (1992).
[3] T. Sato, T. Sueyoshi, and M. Iwatsuku, Surf. Sci. Lett. **321**, L137 (1994).
[4] D. Chen and D. Sarid, Surf. Sci. **319**, 74 (1994).
[5] P. H. Beton, A. W. Dunn, and P. Moriarty, Surf. Sci. **361/362**, 878 (1996).
[6] T. Hashizume, K. Motai, X. D. Wang *et al.*, J. Vac. Sci. Technol. A **12**, 2097 (1994).
[7] T. Chen, S. Howelles, M. Gallager *et al.*, J. Vac. Sci. Technol. B **9**, 2461 (1991).
[8] D. K. Kin, Y. D. Suh, K. H. Park *et al.*, J. Vac. Sci. Technol. A **11**, 1675 (1993).
[9] G. K. Wertheim, Solid State Commun. **88**, 97 (1993).
[10] E. V. Rut'kov, A. Ya. Tontegode, and Yu. S. Grushko, JETP Lett. **57**, 724 (1993)].
[11] E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Phys. Rev. Lett. **74**, 758 (1995).
[12] N. R. Gall, E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Mol. Mater. **7**, 187 (1996).
[13] N. R. Gall', E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Pis'ma Zh. Tekh. Fiz. **23**(23), 56 (1997) [Tech. Phys. Lett. **23**(12), 911 (1997)].
[14] N. R. Gall, E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Appl. Surf. Sci. **78**, 179 (1994).
[15] E. V. Rut'kov, A. Ya. Tontegode, M. M. Usufov, and N. R. Gall', Zh. Tekh. Fiz. **62**(10), 148 (1992) [Sov. Phys. Tech. Phys. **37**, 1038 (1992)].
[16] N. R. Gall, S. N. Mikhailov, E. V. Rut'kov, and A. Ya. Tontegode, Surf. Sci. **191**, 185 (1987).
[17] C. Gripon, L. Legrand, I. Rosenman, and F. Boue, Fullerene Sci. Technol. **4**, 1195 (1996).

Translated by P. Shelnitz

Self-burial of radioactive waste

L. Ya. Kosachevskiĭ and L. S. Syui

Moscow State University of Environmental Engineering, 127550 Moscow, Russia
(Submitted March 24, 1998; resubmitted January 11, 1999)
Zh. Tekh. Fiz. **69**, 123–127 (November 1999)

The problem of the ''self-burial'' of radioactive waste into melting rock is solved for a spherical container of finite thickness. The mathematical model constructed, unlike the existing ones, takes into account the thermal losses to the solid rock and to the melt behind the container, as well as the reverse evolution of heat upon solidification of the melt. A calculation for the particular case of self-burial in granite shows that consideration of these factors significantly increases the maximum permissible radius at which the container will remain in the solid state and slows the burial rate. © *1999 American Institute of Physics.* [S1063-7842(99)02211-4]

One of the promising methods for the final disposal of radioactive waste is ''self-burial.'' Due to the evolution of heat accompanying radioactive decay, a container with radioactive waste melts the surrounding rock and sinks into it under the action of its own weight. The increase in the amount of heat expended on melting the rock during the accelerated downward migration in the initial stage ensures that the process will pass to a steady state. The problem of the steady migration of a spherical heat source was treated in Ref. 1 under the assumption that the temperature of its surface is uniform. As was shown in Ref. 2, this condition does not correspond to reality. The temperature reaches a minimum at the lower critical point and a maximum at the upper, diametrically opposite point. It also increases with increasing radius. Therefore, to keep a container in the solid state, its radius must not exceed a maximum permissible value at which the surface temperature reaches the melting point of the container. The dependence of the limiting radius and the corresponding maximum burial rate on the thickness and thermal conductivity of the container was investigated in Ref. 3. It was assumed in those studies that the heat flux in the direction opposite to the direction of motion can be neglected and that the heat flux in the direction of motion is completely expended on melting the medium. The reverse evolution of heat upon solidification of the melt behind the heat source was not taken into account. The purpose of the present work is to solve the problem of the self-burial of radioactive waste in a spherical container of finite thickness without these assumptions.

The stationary axisymmetric distributions of the temperature in the radioactive waste (T^i) and in the container wall (T^c) satisfy the equations

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial T^i}{\partial r}\right)+\frac{\partial}{\partial \xi}\left[(1-\xi^2)\frac{\partial T^i}{\partial \xi}\right]=-\frac{q}{k_i}r^2, \quad r<R_i,$$

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial T^c}{\partial r}\right)+\frac{\partial}{\partial \xi}\left[(1-\xi^2)\frac{\partial T^c}{\partial \xi}\right]=0, \quad R_i<r<R. \tag{1}$$

Their solution with allowance for the conditions of continuity of the temperature and the heat flux at the inner surface of the wall can be written in the form[3]

$$T^i(r,\xi)=T_m+\frac{4}{3}\frac{h\nu}{c_p}\left[\frac{1}{8}\left(1-\frac{r^{*2}}{\zeta^2}\right)+\frac{k_i}{4k_c}(1-\zeta)+\sum_{n=0}^{\infty}\frac{C_n}{\Omega_n}r^{*n}P_n(\xi)\right],$$

$$T^c(r,\xi)=T_m+\frac{4}{3}\frac{h\nu}{c_p}\left\{\frac{k_i\zeta}{4k_c}\left(\frac{1}{r^*}-1\right)+\sum_{n=0}^{\infty}\frac{C_n}{(2n+1)\Omega_n}\left[\left(n+1+n\frac{k_i}{k_c}\right)r^{*n}+n\left(1-\frac{k_i}{k_c}\right)\zeta^{2n+1}r^{*-n-1}\right]P_n(\xi)\right\}. \tag{2}$$

Hence follow the expression for Stefan's number

$$S(\xi)=\frac{c_p}{h}[T_w(\xi)-T_m]=\frac{4}{3}\nu\sum_{n=0}^{\infty}C_nP_n(\xi),$$

$$T_w(\xi)=T^c(R,\xi) \tag{3}$$

and the expression for the heat flux from the source to the surrounding medium

$$-k_c\frac{\partial T^c}{\partial r}(R,\xi)=\frac{4}{3}\frac{h\nu k_i}{c_pR}\left[\frac{\zeta}{4}-\sum_{n=1}^{\infty}n\Gamma_nC_nP_n(\xi)\right]. \tag{4}$$

Here

$$r^*=\frac{r}{R}, \quad \xi=\cos\Theta, \quad \nu=\frac{c_pqR_i^2}{hk_i}, \quad h=h_m+c_pT_m,$$

$$\Gamma_n=\frac{1}{\Omega_n}\left[1+\frac{n+1}{2n+1}\left(\frac{k_c}{k_i}-1\right)(1-\zeta^{2n+1})\right],$$

$$\Omega_n=1-\frac{n}{2n+1}\left(1-\frac{k_i}{k_c}\right)(1-\zeta^{2n+1}), \quad \zeta=\frac{R_i}{R},$$

C_n are arbitrary constants; r and Θ are spherical coordinates; $P_n(\xi)$ are Legendre polynomials; q is the heat output power of the radioactive waste; k_i and k_c are the thermal conductivities of the radioactive waste and the container; c_p, T_m,

and h_m are the specific heat, melting point, and heat of the phase transition of the medium; R_i and R are the inner and outer radii of the container; and their ratio ζ is chosen from considerations of mechanical strength and is henceforth considered fixed for different values of R.

The region of the melt in front of the heat source ($\xi>0$) forms a thin layer, in which flow is described by the methods of lubrication theory. In the reference frame associated with the source the velocity field and the pressure are specified by the equations[2]

$$v_\Theta = V\frac{y}{\delta^*}\left[1+\frac{3}{\delta^*}\left(1-\frac{y}{\delta^*}\right)\right]\sin\Theta,$$

$$v_r = -V\left(\frac{y}{\delta^*}\right)^2\left\{\left(3-2\frac{y}{\delta^*}\right)\cos\Theta + \frac{d\delta^*}{d\Theta}\left[\frac{1}{2}+\frac{3}{\delta^*}\left(1-\frac{y}{\delta^*}\right)\right]\sin\Theta\right\},$$

$$p = p_0 + 6\frac{\eta V}{R}\int_\Theta^{\pi/2}\frac{\sin\Theta}{\delta^{*3}}d\Theta, \quad y=r^*-1, \quad \delta^*=\frac{\delta}{R}, \tag{5}$$

where δ is the thickness of the layer, η is the viscosity coefficient, and V is the burial rate.

In the region of the melt behind the source ($\xi<0$) the burial rate and pressure are assumed to be constant and equal to V and p_0, respectively. The tangential stresses on the surface of the source are small compared with the pressure. Therefore, the drag force of the melt equals

$$F = 2\pi R^2\int_{-1}^{1} p\xi\, d\xi = 6\pi\eta R V J,$$

$$J = \int_0^1 (1-\xi^2)\,\delta^{*-3}\,d\xi. \tag{6}$$

Equating the difference between the weight and the buoyant force to this expression for the drag, we obtain the equation

$$VJ = \frac{2}{9}\frac{g}{\eta}R^2(\rho_1-\rho), \quad \rho_1 = \rho_c + \zeta^3(\rho_i-\rho_c), \tag{7}$$

where ρ_i, ρ_c, and ρ are the densities of the radioactive waste, the container material, and the medium, and g is the acceleration of gravity.

The main contribution to the integral J is made by a small vicinity about the point $\xi=1$. Now setting

$$\delta^*(\xi) = \delta^*(1) - \delta^{*\prime}(1)(1-\xi)$$

and taking δ^* as the integration variable, we obtain up to the leading term

$$J = \frac{1}{\delta^*(1)[\delta^{*\prime}(1)]^2}. \tag{8}$$

The axisymmetric distribution of the temperature in the solid phase T^s satisfies the equation

$$-V\left(\xi\frac{\partial T}{\partial r}+\frac{1-\xi^2}{r}\frac{\partial T}{\partial \xi}\right) = \frac{a}{r^2}\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right)+\frac{\partial}{\partial\xi}\left[(1-\xi^2)\frac{\partial T}{\partial\xi}\right]\right\}. \tag{9}$$

According to the assumption made above, this equation is also valid for the distribution of the temperature T^f in the melt at $\xi<0$. Its solution has the form

$$T^s(r,\xi) = \frac{1}{\sqrt{r^*}}\exp\left(-\frac{\beta}{2}r^*\xi\right) \times\sum_{n=0}^{\infty} E_n K_{n+\frac{1}{2}}\left(\frac{\beta}{2}r^*\right)P_n(\xi),$$

$r>R+\delta$,

$$T^f(r,\xi) = T_m + \frac{1}{\sqrt{r^*}}\exp\left(-\frac{\beta}{2}r^*\xi\right)\sum_{n=0}^{\infty}\left[F_n K_{n+\frac{1}{2}}\left(\frac{\beta}{2}r^*\right) + G_n I_{n+\frac{1}{2}}\left(\frac{\beta}{2}r^*\right)\right]P_n(\xi),$$

$R<r<R+\delta, \quad \xi<0,$ (10)

where $K_{n+\frac{1}{2}}(x)$ and $I_{n+\frac{1}{2}}(x)$ are Bessel functions of imaginary argument, $\beta = VR/\mathfrak{a}$, and $\mathfrak{a} = k/\rho c_p$.

The density, specific heat, and thermal conductivity of the medium in the solid and liquid phases are assumed to be identical. The arbitrary constants E_n, F_n, and G_n are specified by the boundary conditions

$$T^s(R+\delta,\xi) = T^f(R+\delta,\xi) = T_m,$$

$$T^f(R,\xi) = T^c(R,\xi). \tag{11}$$

Confining ourselves to the leading terms of the asymptotic expressions of the Bessel functions

$$K_{n+\frac{1}{2}}(x) = \left(\frac{\pi}{2x}\right)^{1/2}\exp(-x),$$

$$I_{n+\frac{1}{2}}(x) = \frac{1}{(2\pi x)^{1/2}}[\exp(x)-(-1)^n\exp(-x)],$$

we have

$$T^s(r,\xi) = T_m\frac{1+\delta^*}{r^*}\exp\left[-\frac{\beta}{2}(r^*-1-\delta^*)(1+\xi)\right],$$

$r>R+\delta$,

$$T^f(r,\xi) = T_m + \frac{h}{c_p}\frac{S(\xi)}{r^*}\frac{1-\exp[-\beta(1+\delta^*-r^*)]}{1-\exp(-\beta\delta^*)} \times\exp\left[-\frac{\beta}{2}(r^*-1)(1+\xi)\right],$$

$R<r<R+\delta, \quad \xi<0.$ (12)

From the boundary condition

$$-k\frac{\partial T^f}{\partial r}(R+\delta,\xi)=\pm h_m\rho V\xi-k\frac{\partial T^s}{\partial r}(R+\delta,\xi), \qquad (13)$$

where $\pm h_m\rho V\xi$ is the quantity of heat spent on melting the medium in front of the heat source and given back upon solidification behind it, we obtain the equation of the phase boundary at $\xi<0$

$$(1+\delta^*)[D+(3D-1)\xi]+\frac{2D}{\beta}$$
$$=\frac{S(\xi)}{1-\exp(-\beta\delta^*)}\exp\left[-\frac{\beta}{2}\delta^*(1+\xi)\right],$$
$$D=\frac{c_pT_m}{2h} \qquad (14)$$

and the expression for the heat flux from the melt into the solid phase at $\xi>0$

$$-k\frac{\partial T^f}{\partial r}(R+\delta,\xi)=\frac{kh}{c_pR}[(1-D)\beta\xi+D(\beta+2)]. \qquad (15)$$

The distribution of the temperature in the melt in front of the heat source is found by a parametric method of boundary-layer theory. Integrating the heat conduction equation

$$v_r\frac{\partial T^f}{\partial y}+v_\Theta\frac{\partial T^f}{\partial\Theta}=\frac{\mathfrak{a}}{R}\frac{\partial^2T^f}{\partial y^2}$$

over the thickness of the layer and taking into account the continuity equation

$$\sin\Theta\frac{\partial v_r}{\partial y}+\frac{\partial}{\partial\Theta}(v_\Theta\sin\Theta)=0,$$

we obtain the integral relation

$$\left[\int_0^{\delta^*}v_\Theta T^f\sin\Theta\,dy\right]'+VT_m[\xi-(1-\xi^2)\delta^{*\prime}]$$
$$=-\frac{a}{R}\left[\frac{\partial T^f}{\partial y}(\delta,\xi)-\frac{\partial T^f}{\partial y}(0,\xi)\right]. \qquad (16)$$

We approximate T^f by a trinomial, which is quadratic with respect to y and whose coefficients are defined by the conditions (11) and (15),

$$T^f(y,\xi)=T_m+\frac{h}{c_p}\left(1-\frac{y}{\delta^*}\right)$$
$$\times\left\{S+y\left[D(\beta+2)+(1-D)\beta\xi-\frac{S}{\delta^*}\right]\right\}. \qquad (17)$$

The relation (16) with allowance for (5) and (17) gives the differential equation for δ

$$\frac{1-\xi^2}{20}\left[D(\beta-8)+(1-D)\beta\xi+\frac{5}{6}S\right]\delta^{*\prime}=\varphi-\frac{S}{\beta\delta^*},$$

$$\varphi=D\left(1+\frac{2}{\beta}\right)+(1-D)\xi-\frac{3}{40}[(1-\xi^2)S]', \qquad (18)$$

whence it follows that up to the leading term we have

$$\delta^*=\frac{S}{\beta\varphi}. \qquad (19)$$

The prime sign denotes a derivative with respect to ξ. Using (8) and (19), we bring Eq. (7) into the form

$$S(1)=\frac{9}{2}\frac{\gamma\beta(\beta+2D)^3}{[S'(1)]},\quad \gamma=\frac{\mathfrak{a}\eta}{gR^3(\rho_1-\rho)}. \qquad (20)$$

From (12) and (14) we find

$$-k\frac{\partial T^f}{\partial r}(R,\xi)$$
$$=\begin{cases}\dfrac{kh\beta}{c_pR}\left\{D\left(1+\dfrac{2}{\beta}\right)\right.\\ \quad\left.+(1-D)\xi-\dfrac{3}{20}[(1-\xi^2)S]'\right\}, & \xi\geqslant 0\\ \dfrac{kh\beta}{c_pR}\left\{\left(\dfrac{1+\xi}{2}+\dfrac{1}{\beta}\right)S\right.\\ \quad+\left[(1+\delta^*)(D+3D\xi-\xi)+\dfrac{2D}{\beta}\right]\\ \quad\left.\times\exp\left[-\dfrac{\beta}{2}\delta^*(1-\xi)\right]\right\}, & \xi<0.\end{cases} \qquad (21)$$

To determine the constants C_n we have the boundary condition

$$k_c\frac{\partial T^c}{\partial r}(R,\xi)=k\frac{\partial T^f}{\partial r}(R\xi). \qquad (22)$$

In the limiting case of $\beta=0$ and $\delta=0$, it follows from (14) and (19) that $S=0$ and $C_n=0$, i.e., we obtain the stationary solution for an immobile container with a surface temperature T_m. The heat flux (21) in all directions reduces to kT_m/R. According to (4) and (22), the radius of such a container equals

$$R_0=\left(\frac{3kT_m}{q\zeta^3}\right)^{1/2}. \qquad (23)$$

Thus, self-burial is possible under the condition $R>R_0$. As R is increased, the exponential functions in (21) rapidly decrease; therefore, we shall henceforth neglect the corresponding terms. The resulting error in the determination of the burial rate is not more than V_0, i.e., the value obtained for $R=R_0$. Multiplying (22) by $P_n(\xi)$ and integrating over ξ from -1 to $+1$, we obtain the infinite system of equations

$$\frac{4n\chi}{(2n+1)\beta}C_n+\sum_{k=0}^{\infty}C_k\left\{\frac{3}{10}P_k(0)P_n(0)+\frac{n}{2n+1}\left[\frac{3}{10}(n+1)-(-1)^{k+n}\right]\Psi_{k,n-1}-\frac{n+1}{2n+1}\left[\frac{3}{10}n+(-1)^{k+n}\right]\Psi_{k,n+1}+(-1)^{k+n}\left(1+\frac{2}{\beta}\right)\Psi_{k,n}\right\}=\frac{\chi}{2}\frac{\zeta}{\beta}\int_{-1}^{1}P_n(\xi)\,d\xi-\frac{3}{2\nu}Q_n,$$

$$Q_n=D\left(1+\frac{2}{\beta}\right)\Psi_{0,n}+(1-D)\Psi_{1,n},$$

$$\Psi_{k,n}=\int_0^1 P_k(\xi)P_n(\xi)\,d\xi,\quad \chi=\frac{k_i}{k},\quad n=0,1,\ldots. \tag{24}$$

We confine ourselves to a finite number N of the constants C_n and the first N equations. For $N=2$ we have

$$C_0=\frac{3}{\nu}\frac{\Delta_0}{\Delta},\quad C_1=-\frac{3}{\nu}\frac{\Delta_1}{\Delta};$$

$$S(1)=\frac{4}{3}\nu(C_0+C_1),\quad S'(1)=\frac{4}{3}\nu C_1;$$

$$\Delta_0=\frac{2}{3}\chi\zeta\nu\left(1+2\chi\Gamma_1+\frac{19}{80}\beta\right)-\frac{1}{2}(7+8\chi\Gamma_1)D-\left[1+\frac{59}{40}D+\chi\Gamma_1(1+D)\right]\beta-\frac{1}{480}(97+77D)\beta^2,$$

$$\Delta_1=\chi\zeta\nu\left(-1+\frac{1}{30}\beta\right)+6D+\frac{1}{20}(35+47D)\beta+\frac{1}{40}(15+7D)\beta^2,$$

$$\Delta=1+8\chi\Gamma_1+\frac{1}{20}(43+64\chi\Gamma_1)\beta+\frac{119}{300}\beta^2. \tag{25}$$

The parameter β is found from the equation

$$\Delta_0-\Delta_1=\frac{\Delta}{4}S(1), \tag{26}$$

where, according to (20),

$$S(1)=\frac{9\chi}{32}\left(\frac{\Delta}{\Delta_1}\right)^2\beta(\beta+2D)^3. \tag{27}$$

Since $S(1)\ll 1$, Eq. (26) reduces to a quadratic equation. When $R>R_0$, it has one positive root

$$\beta(\nu)=\frac{1}{277+161D}\{[36L^2+(277+161D)M]^{1/2}-6L\},$$

$$L=110+153D+40\chi\Gamma_1(1+D)-5\chi\zeta\nu,$$

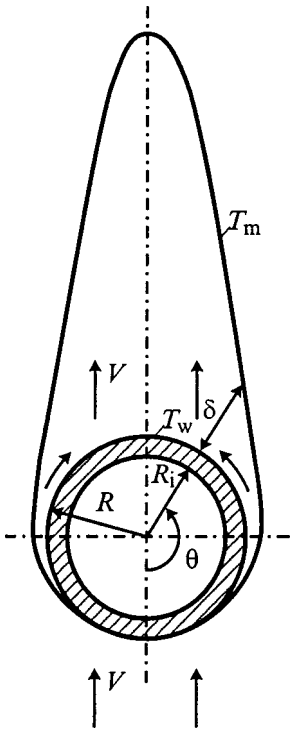

FIG. 1. Physical model of the geometry of the problem.

$$M=\frac{1}{3}\chi\zeta\nu(5+4\chi\Gamma_1)-\frac{1}{2}(19+8\chi\Gamma_1)D, \tag{28}$$

which specifies the dependence of the burial rate on the radius and heat output power of the heat source, as well as the physical characteristics of the medium. According to (2), the maximum temperature is achieved on the line $\xi=-1$ at the point

$$r_m^*=-4\frac{\zeta^2}{\Omega_1}C_1; \tag{29}$$

and equals

$$T^i(r_m^*,-1)=T_m+\frac{h\nu}{6c_p}\left[1+2\frac{k_i}{k_c}(1-\zeta)+\frac{r_m^*}{\zeta}(r_m^*+2\Omega_1)\right]. \tag{30}$$

The container has its highest temperature at the upper critical point on the inner surface of the wall:

$$T^c(R_i,-1)=T_m+\frac{4}{3}\frac{h\nu}{c_p}\left[\frac{k_i}{4k_c}(1-\zeta)+\left(1+\frac{\zeta}{\Omega_1}\right)C_0\right]. \tag{31}$$

Equating it to the melting point of the container T_*, we find the maximum permissible value of ν:

$$\nu_*=\frac{3}{G}\left\{\left(1+\frac{\zeta}{\Omega_1}\right)[876D+(277+161D)\beta_*]-S_*(445+464\chi\Gamma_1)\right\},$$

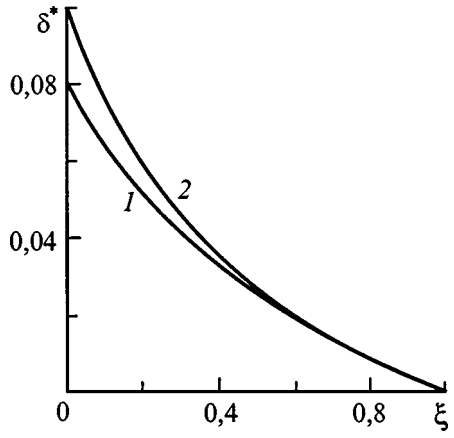

FIG. 2. Thickness of the melt zone at $\xi>0$.

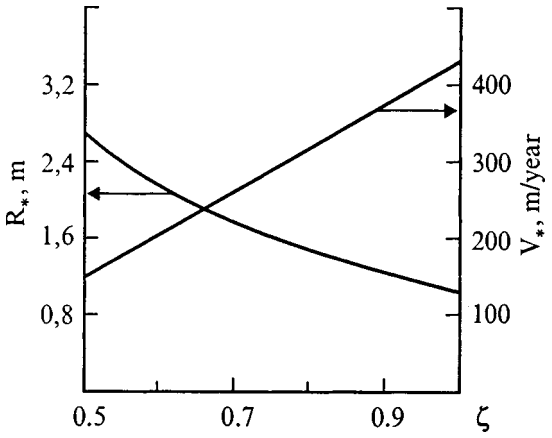

FIG. 4. Plots of the dependence of the maximum radius and burial rate on ζ.

$$G=180\chi\zeta\left(1+\frac{\zeta}{\Omega_1}\right)-\frac{k_i}{k_c}(445+464\chi\Gamma_1)(1-\zeta),$$

$$S_*=\frac{c_p}{h}(T_*-T_m). \tag{32}$$

Substituting it into (26), we obtain a quadratic equation for the corresponding value $\beta=\beta_*$. According to the definitions of these parameters, the maximum radius and burial rate are found using the formulas

$$R_*=\frac{1}{\zeta}\left(\frac{hk_i}{c_pq}\nu_*\right)^{1/2}, \quad V_*=\frac{\alpha}{R_*}\beta_*. \tag{33}$$

When $N>2$, the problem requires a numerical solution.

Let us consider the case of the self-burial of radioactive waste housed in a container composed of the high-temperature ceramic NbC in granite when $q=130\,000$ W/m^3. We take the following values for the physical constants (in SI units):[4] $\rho=2700$, $c_p=1301$, $k=3.013$, $h_m=585\,800$, $T_m=1200\,^\circ$C, $\eta=10$; $\rho_c=7820$, $k_c=44$, $T_*=3480\,^\circ$C; $\rho_i=7800$, $k_i=36$.

For $\zeta=0.9$ we obtain $R_*=1.221$ m, $V_*=376.28$ m/year, $S(1)=0.855\times10^{-4}$, $S'(1)=-0.684$, $\delta^*(1)=0.359\times10^{-5}$, and $\delta^*(-1)=4.012$.

At the point $\xi=0$ the expressions (14) and (19) give fairly close values: $\delta^*(-0)=0.092$ and $\delta^*(+0)=0.099$. At $\xi=1$ the container is in direct contact with the solid medium and has a temperature exceeding its melting point by 0.14 °C. The maximum temperature within the container at the point $r_m^*=0.639$, $\xi=-1$ equals 3540.96 °C. The melt zone and the computational model of flow are shown in Fig. 1. In Fig. 2 curve *1* corresponds to the boundary of the melt zone in the region $\xi>0$ defined by Eq. (18), and curve *2* corresponds to the approximate formula (19). In the range $0.5<\xi\leqslant1$ they essentially coincide. The maximum difference between them at $\xi=0$ amounts to 0.019. Since the motion of the container depends mainly on the conditions in the vicinity of $\xi=1$, the accuracy of formula (19) is fully satisfactory. Formula (6) with allowance for (19) gives $J=139.04\times10^6$, and the value obtained from the approximate formula (8) is 0.05% higher. According to (23), $R_0=0.338$ m corresponds to a zero burial rate. Figure 3 shows a plot of $V(R)$ specified by formula (28). For the radius R_0 it gives $V_0=68.39$ m/year, which represents the maximum error caused by neglect of the exponential term in (21). As ζ is increased from 0.5 to 1, the radius R_* decreases from 2.683 to 1.025 m, and V_* increases roughly according to a linear law from 150.08 to 429.94 m/year (Fig. 4). A comparison with the results in Ref. 2 for $\zeta=1$ shows that consideration of the heat flux and the reverse evolution of heat upon solidification of the melt at $\xi<0$ increases R_* by 0.298 m and diminishes V_* by 44 m/year. The melt zone behind the container becomes 1.5 times longer.

The results obtained depend weakly on the choice of the value of N. For example, for $N=10$ the values of R_* and V_* increase by 1.53% and 0.51%, respectively, and for $N=100$ they increase by 1.62% and 0.56%.

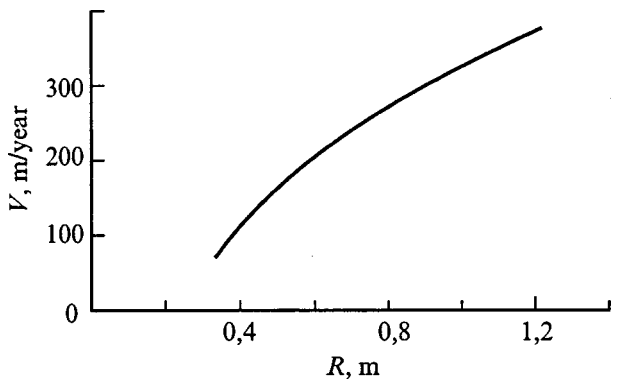

FIG. 3. Plot of the dependence of the burial rate on the container radius for $\zeta=0.9$.

[1] S. H. Emerman and D. L. Turcotte, Int. J. Heat Mass Transf. **26**, 1625 (1983).
[2] L. Ya. Kosachevskiĭ and L. S. Syui, Zh. Tekh. Fiz. **64**(6), 7 (1994) [Tech. Phys. **39**, 524 (1994)].
[3] L. Ya. Kosachevskiĭ and L. S. Syui, Geoékologiya, No. 1, pp. 96–102 (1995).
[4] I. K. Kikoin (ed.), *Tables of Physical Constants. A Handbook* [in Russian], Atomizdat, Moscow (1976), 1008 pp.

Translated by P. Shelnitz

BRIEF COMMUNICATIONS

Interaction of Ag with faceted $Pb_{1-x}Sn_xTe$ single crystals

M. V. Bestaev, V. A. Moshnikov, and A. I. Rumyantseva

St. Petersburg Electrical-Engineering University, 197376 St. Petersburg, Russia
(Submitted June 3, 1997)
Zh. Tekh. Fiz. **69**, 128–129 (November 1999)

The results of an investigation of the interaction of silver atoms with $Pb_{1-x}Sn_xTe$ single crystals having natural faceting are presented. A model is proposed for the mass transport of low-volatility dopants through the vapor phase in the form of tellurides. © *1999 American Institute of Physics.* [S1063-7842(99)02311-9]

Semiconducting solid solutions between lead telluride and tin telluride are widely used in IR optoelectronics for injection lasers and photodetectors.[1] Numerous studies have been devoted to aspects of their doping. The behavior of many impurities has a complex character. It was found in Ref. 2 for the doping of PbTe with tin that layers of semiconducting solid solutions form on the surface from the vapor phase and that the diffusion parameters depend significantly on the concentration of intrinsic defects. When PbTe and $Pb_{1-x}Sn_xTe$ are doped with gallium, microinclusions of complex character with compositions corresponding to the formulas $PbGa_6Te_{10}$ and $Pb_{1-x}Sn_xGa_6Te_{10}$ can appear.[3] The mechanism of mass transport through the vapor phase is satisfactorily described by model representations of the interaction of low-volatility dopants with tellurium vapor. Such components are transported in the form of tellurides, particularly the tellurides of tin, indium, germanium, and gallium.[4]

When PbTe and $Pb_{1-x}Sn_xTe$ are doped with zinc and cadmium, the mass transport has a fundamentally different character. The tellurides of zinc and cadmium exhibit a high degree of dissociation in the vapor phase,[5] with high values for the partial pressures of zinc and cadmium. As a result, a reaction of the type[6]

$$Zn + PbTe \rightarrow ZnTe + Pb$$

takes place on the surface of the crystal.

When the structural quality of the crystals being treated is high, a multilayer ZnTe/Pb/PbTe structure can be created.[7] This provides direct experimental evidence of the need to take into account the counterflow of tellurium atoms when an impurity diffuses in lead telluride.

The complex behavior of silver atoms in lead telluride and solid solutions based on it was noted in Ref. 8. The admixture of Ag to the original mixture led to an increase in the concentration of charge carriers by tens of times in comparison to the concentration of the impurity introduced.

This paper presents the results of an investigation of the interaction of silver atoms with $Pb_{1-x}Sn_xTe$ single crystals having natural faceting.

The $Pb_{1-x}Sn_xTe$ single crystals were grown from the vapor phase using the method described in Ref. 7. The crystals had crystallographic (100) faceting and *p*-type conduction with a carrier concentration equal to $(2-5)\times 10^{18}\,cm^{-3}$, a dislocation density equal to $10^4\,cm^{-2}$, and faces measuring 3×4 mm.

It follows from Refs. 2–7 that the possibility of an impurity reaching the surface of a crystal in ''pure'' form or in the composition of a telluride must be taken into account in investigations of diffusion processes. In such cases the character of the interaction will be different for each of the components, and the ratio between the components is determined by the thermodynamic and kinetic conditions under which the experiment is carried out.[9] For model isolation of the role of ''pure'' silver, Ag was preliminarily deposited on a face of an unheated crystal.

The subsequent diffusion process was carried out in an evacuated ampul under isothermal conditions. In order to avoid direct contact between the dopant and the crystal, the latter was placed in a quartz flask. The reaction took place through the vapor phase, the annealing temperature was varied in the range 773–973 K, and the isothermal annealing time was $0.5-5$ h. The distribution of elements along the direction of diffusion, which was perpendicular to a surface of natural (100) faceting, was analyzed using x-ray microanalysis. A spread of values of the mass concentrations over the surface of the face of the $Pb_{1-x}Sn_xTe$ single crystals was established. The formation of tiny needles, which attest to growth according to a vapor/liquid/crystal mechanism, was noted in several experiments. In all cases islands containing Ag_2Te formed in the near-surface layers. All the results observed can be attributed to the occurrence of the following reaction on the crystal surface:

$$Pb_{1-x}Sn_xTe + 2Ag \rightarrow Ag_2Te + xSn + (1-x)Pb.$$

The formation of free phases of Sn and Pb can lead to the appearance of microdroplets followed by the growth of acicular crystals. Thus, when silver diffuses from the vapor phase, mass transport can occur in the crystal in two directions: silver can diffuse into the bulk of the sample, and tellurium can diffuse to the surface. In addition, the excess metallic components can separate as free phases. All three

processes promote sharper variation of the electrophysical properties of $Pb_{1-x}Sn_xTe$ single crystals. Such processes can also occur in bulk doped tin–lead telluride crystals when supersaturated solid solutions decompose.

[1] N. Kh. Abrikosov and L. E. Shelimova, *Semiconductor Materials Based on IV–VI Compounds* [in Russian], Nauka, Moscow (1975), 196 pp.
[2] M. V. Bestaev, T. T. Dedegkaev, and V. A. Moshnikov, Fiz. Tverd. Tela (Leningrad) **26**, 2200 (1984) [Sov. Phys. Solid State **26**, 1335 (1984)].
[3] Sh. M. Duguzhev, A. V. Makhin, V. A. Moshnikov, and A. L. Shelykh, Cryst. Res. Technol. **25**(2), 145 (1990).
[4] M. V. Bestaev, T. T. Dedegkaev, T. B. Zhukova *et al.*, *Abstracts of the 1st Ural Conference ''Surface and New Materials''* [in Russian], Sverdlovsk (1984), p. 177.
[5] M. Aven and J. S. Prener (Eds.), *Physics and Chemistry of II–VI Compounds* [North-Holland, Amsterdam (1967); Mir, Moscow (1970), 624 pp.].
[6] A. V. Novoselova, V. P. Zlomanov, A. M. Gas'kov *et al.*, Vestn. Mosk. Univ., Ser. 2: Khim. **22**, 3 (1982).
[7] M. V. Bestaev, K. Gatsoev, A. I. Gorelik, and V. A. Moshnikov, *International Conference ''Material Science and Material Properties for Infrared Optoelectronics,''* Uzhgorod, Ukraine (1996).
[8] G. A. Kalyuzhnaya and K. V. Kiseleva, in *Stoichiometry in Crystal Compounds and Its Influence on Their Physical Properties (Proceedings of the Lebedev Physics Institute, Academy of Sciences of the USSR, Vol. 177)*, edited by Yu. M. Popov [Nova Science, Commack, N.Y. (1988); Tr. Fiz. Inst. Akad. Nauk SSSR **177**, 5–84 (1987)].
[9] V. P. Zlomanov and A. V. Novoselova, *$P-T-X$ Phase Diagrams of Metal–Chalcogen Systems* [in Russian], Nauka, Moscow (1987), 208 pp.

Translated by P. Shelnitz

Potential electrical characteristics of interference transistors made from various materials

I. I. Abramov, Yu. A. Berashevich, and A. L. Danilyuk

Belarus State University of Informatics and Radio Electronics, 220027 Minsk, Belarus
(Submitted October 19, 1998)
Zh. Tekh. Fiz. **69**, 130–131 (November 1999)

A theoretical analysis of the electrical characteristics of GaAs, InAs, InSb, and Si quantum interference *T* transistors is performed with consideration of the dependence of the effective masses on the quantum wire dimensions. It is shown for extremely small wire dimensions that none of the materials has significant advantages over the others with respect to the frequency characteristics of the transistors investigated. © *1999 American Institute of Physics.* [S1063-7842(99)02411-3]

Along with the familiar single-electron tunneling and resonant-tunneling transistors, quantum wire structures are also promising for creating various high-speed electronic and optoelectronic devices.[1] One such structure, which is of considerable interest, is the quantum interference *T* transistor.[2,3] For example, a theoretical analysis of GaAs *T* transistors (see Fig. 1) was carried out in Ref. 4, and their merits and drawbacks were demonstrated.

The purpose of the present work is to ascertain the potential electrical characteristics of *T* transistors made from various materials by carrying out a theoretical investigation.

For this purpose, transistors made from four promising nanoelectronics materials,[5] viz., GaAs, InAs, InSb, and Si, were analyzed in the present study. Very small geometric dimensions were chosen for the investigation. The lengths of the devices were $d_2 = 1000$ Å, and their widths were $d_1 = 10$ Å. The cross-sectional area of the wires was $S = 10 \times 10$ Å. An extremely small quantum wire width was chosen owing to the significant progress that has been made in nanotechnology methods and the already available possibility of fabricating wires with a width of about 20 Å (Ref. 6). The analysis was performed at $T = 4.2$ K in order to minimize the influence of scattering processes.

The model of interference *T* transistors is based on the scattering-matrix formalism.[7] The fundamental difference between the model of the device that we used and the model in Ref. 4 was the inclusion of the dependence of the effective mass on the quantum wire dimensions due to their extremely small values. The model described in Ref. 8 was used for these purposes.

The geometric dimensions and modification of the model of a *T* transistor indicated led to the following highly significant differences between our results and the investigations in Ref. 4. It was found that consideration of the dependence of the effective masses on dimensions leads to the appearance of a saturation region on the current–voltage characteristics even at small values of the source–drain voltage V_{sd} less than 50 mV. The results for the devices of the type in Fig. 1a made from various materials are presented in Fig. 2a. It is seen that the characteristics have the typical form for field-effect transistors. Therefore, our conclusion contradicts the conclusion in Ref. 4 that *T* transistors are not applicable to signal amplification because of the absence of saturation at working voltages. We note that the achievement of saturation is highly important not only for analog, but also for logical integrated circuits.

The results of the calculation of the maximum output frequency $f_{\max 1}$ of the transistors of the type in Fig. 1a for various materials with allowance for the dependence of the effective masses on dimensions and without allowance for this dependence are presented in Figs. 2b and 2c, respectively. It is seen that allowance for the dependence of the frequency characteristics on quantum wire width is of fundamental importance for estimating the potential possibilities of a transistor. It was found that for each material there is a carrier concentration n_l (one-dimensional) at which the frequency characteristics are optimal. We also note that the maximum values of the frequencies for different materials differ less significantly than in the case of neglect of depen-

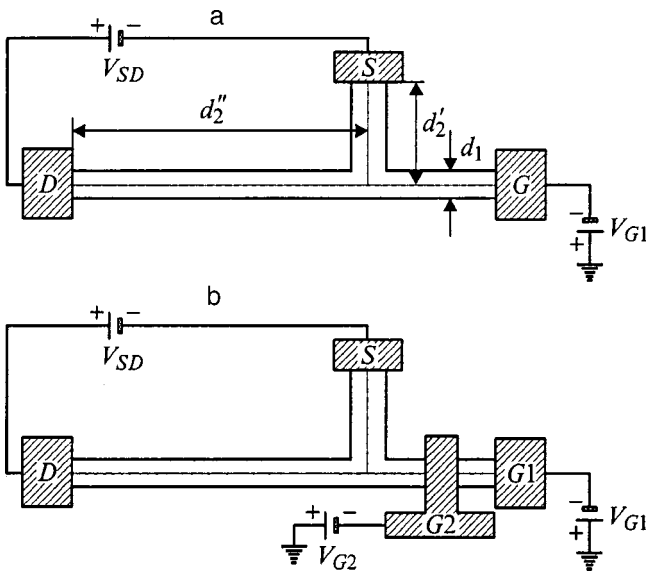

FIG. 1. Structures of *T* transistors: a — single-gate ($d_2 = d_2' + d_2''$), b — double-gate; *S* — source; *D* — drain; *G*, G_1, G_2 — gates.

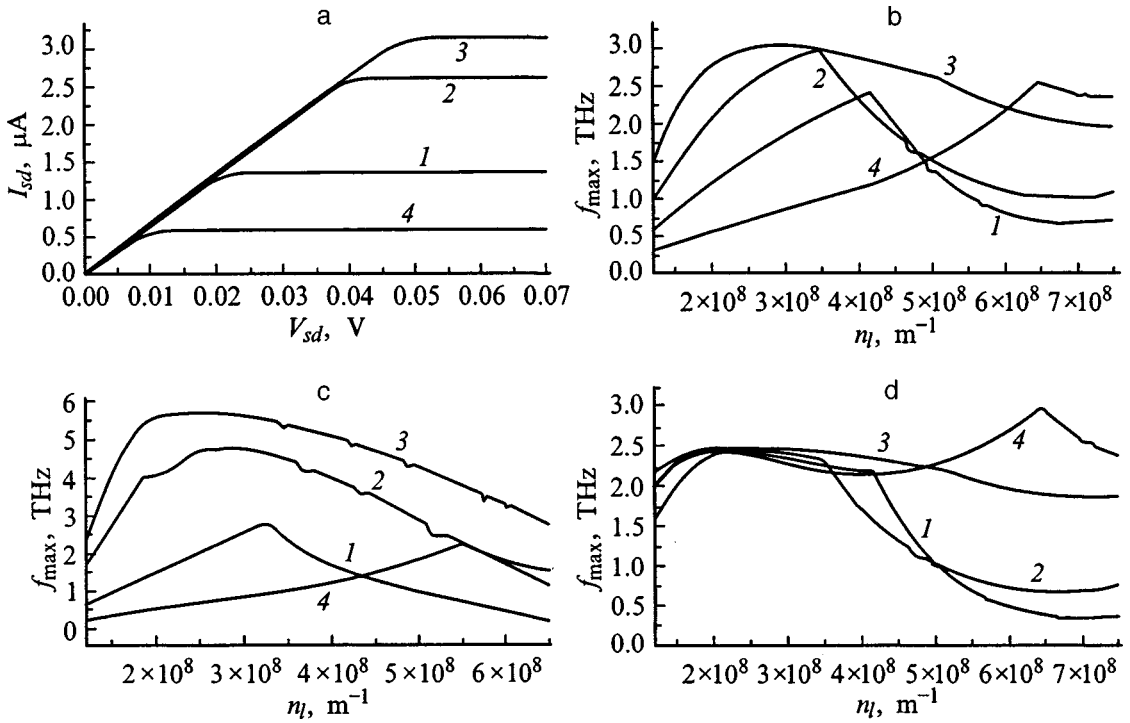

FIG. 2. Electrical characteristics of *T* transistors made from various materials (*1* — GaAs, *2* — InAs, *3* — InSb, *4* — Si): current–voltage characteristics (a) of transistors of the type in Fig. 1a with $n_l = 1.6 \times 10^8$ m^{-1} for $V_g = 1$ V; variation of $f_{\max 1}$ for transistors of the type in Fig. 1a in the cases when the dimensions of the device are taken into consideration (b) and neglected (c); variation of $f_{\max 2}$ for a transistor of the type in Fig. 1b with consideration of the dimensions of the device (d).

dence of the effective masses on dimensions (Fig. 2c). This is associated mainly with the small difference between the increasing values of the masses for all the materials at the quantum wire dimensions under consideration.

However, the most unexpected result was obtained for the maximum output frequency $f_{\max 2}$ (Fig. 2d) of the transistor with two gates (Fig. 1b). As can be seen from the figure, the silicon transistor has acceptable (not poor) frequency characteristics over the entire range of carrier concentrations investigated.

Thus, when the wire widths are extremely small, none of the materials investigated has significant advantages over the others in regard to the frequency characteristics of the quantum interference *T* transistors analyzed. This conclusion is qualitatively consistent with the results in Ref. 9, which were obtained, however, as a result of simulation of the electrical characteristics of MIS transistors as elements of VLSI integrated circuits with extremely small channel lengths.

This work was carried out with partial financial support from the Republic Scientific-Technical Programs ''Informatics,'' ''Low-Dimensional Systems,'' and ''Nanoelectronics.''

[1] Zh. I. Alferov, Fiz. Tekh. Poluprovodn. **32**, 3 (1998) [Semiconductors **32**, 1 (1998)].
[2] S. Datta, Superlattices Microstruct. **6**(1), 83 (1989).
[3] M. A. Reed and W. P. Kirk (eds.), *Nanostructure Physics and Fabrication*, Academic Press, Boston (1989), 517 pp.
[4] S. Subramaniam, S. Bandyopadhyay, and W. Porod, J. Appl. Phys. **68**, 4861 (1990).
[5] I. A. Obukhov, *Proceedings of the 7th International Crimean Microwave Conference* [in Russian], Sevastopol (1997), pp. 383–385.
[6] H. I. Liu, D. K. Biegelsen, F. A. Ponce *et al.*, Appl. Phys. Lett. **64**, 1383 (1994).
[7] B. Shapiro, Phys. Rev. Lett. **50**, 747 (1983).
[8] A. Ghoshal, D. Mitra, and K. P. Ghatak, Nuovo Cimento D **12**, 891 (1988).
[9] M. V. Fishetti and S. E. Laux, IEEE Trans. Electron Devices **ED-38**, 634 (1991).

Translated by P. Shelnitz

Theory of the breakup of a liquid jet into drops

S. K. Aslanov

I. I. Mechnikov Odessa State University, 270026 Odessa, Ukraine
(Submitted October 19, 1998)
Zh. Tekh. Fiz. **69**, 132–133 (November 1999)

A theory for an approximate description of the breakup of a thin liquid jet issuing from an orifice into drops is devised on the basis of mass and energy balance equations. Attention is focused mainly on the effects of the surface tension forces. The mathematical analysis performed permits unequivocal estimation of the relative value of the mean diameter of the drops formed and the distances between them, which agree closely with the known results of experimental observations and the value of the Rayleigh wavelength from the linear theory of jet instability. © *1999 American Institute of Physics.* [S1063-7842(99)02511-8]

A thin liquid jet issuing through a circular orifice under pressure breaks up into drops in the surrounding atmosphere even at low velocities. This phenomenon has found broad application in technology. Neglecting the effects of viscosity, compressibility, and gravity (where possible), Rayleigh[1] attributed this phenomenon to the instability of the cylindrical liquid body formed, which develops mainly from surface tension forces. Their action leads to progressive growth of any random local decrease in the jet diameter relative to its mean value of $2a$, which is accompanied by a corresponding expulsion of liquid and the formation of nearby bulges on the jet. As a result, there is regular alternation of contracted and expanded segments of the initially equilibrium, circular cylindrical configuration of the liquid column. The pinched segments, which gradually elongate, rupture with the formation of tiny droplets, while the bulging segments transform into large (main) drops, which are separated by equal distances and are subjected to deformation pulsations. It was shown in Ref. 1 on the basis of a linear analysis of small axisymmetric disturbances of a jet (disturbances of the form $\sim F(r)\cdot\exp(ikz+\omega t)$, where $\lambda=2\pi/k$) that their maximum growth rate corresponds to the wavelength $\lambda_*=4.51\cdot(2a)$.

In the present work an attempt is made to calculate analytically the process of regular breakup so as to obtain a value of the diameter $2R$ of the drops formed that would agree quantitatively with the known experimental observations[2] and a value of the distance between drops L that would be close to λ_*. Since the hydrodynamic instability of the column of issuing liquid will be most developed at its end, we shall consider the process of detachment of the drop formed in the last of the contraction–expansion cycles repeated along the length of the jet, which is indicated by the shaded area in Fig. 1 (the solid lines schematically represent the equilibrium configuration of the jet, and the dashed lines show its disturbed state). Assuming that such a process is regular, we can apply mass and energy balance equations to this control volume. The liquid is assumed to be incompressible and inviscid, so that for small rates of nearly horizontal flow, our main attention can be concentrated on the effects of the surface tension forces. The appearance of a difference Δl between the length of the contracted segments and the length of the expanded segments (l) should be expected in the nonlinear stage of development of the wave-like axisymmetric deviations of the jet surface from its initial form, which is cylindrical on the average. To within accuracy to the tiny intermediate droplets K formed when the constriction ruptures, the total initial mass of the last cycle of the jet is used to form the main drop of diameter $2R$ breaking off, which is shown in Fig. 2. After this, the jet end AB again takes the originally assumed hemispherical shape under the effect of surface tension. As a result, the mass balance is written in the form

$$(2l+\Delta l)\pi a^2=(4/3)\pi R^3. \tag{1}$$

The energy acquired as a result of the total work performed by the surface tension forces to contract the jet over the entire segment $ABCD$ will be expended on the work for formation of the main drop, i.e., its expansion from the initial

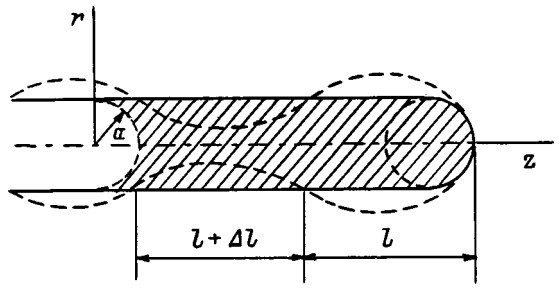

FIG. 1.

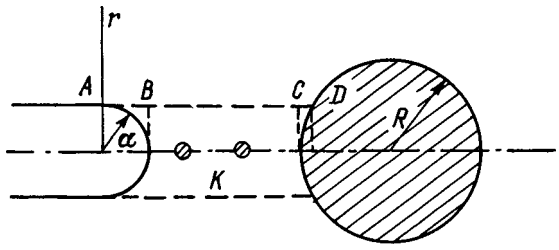

FIG. 2.

diameter $2a$ to the final diameter $2R$ as result of the influx of the displaced liquid. The local surface tension stress $\sigma_r=\sigma/r$ acting on the cylinder of radius r produces a force for the segment BC of complete collapse (rupture) of the jet constriction (Fig. 2). This force is constant over the entire pinched segment, so that the value of the corresponding work can be expressed as $W_1=2\pi\sigma a(l+\Delta l)$. The work which has been performed to expel the liquid in segment AB at the end of the cycle under consideration is expressed by the integral

$$W_2=2\pi\sigma\int_0^a(a-\sqrt{a^2-x^2})\,dx=2\pi\sigma\left(1-\frac{\pi}{4}\right)a^2.$$

In segment CD expulsion of the liquid will occur up to the surface of the drop formed, so that the corresponding work can be represented in a similar manner, i.e.,

$$W_3=2\pi\sigma\int_{-R}^{-\sqrt{R^2-a^2}}(a-\sqrt{R^2-x^2})\,dx$$

$$=2\pi\sigma\left\{a(R-\sqrt{R^2-a^2})\right.$$

$$\left.+\frac{1}{2}\left[a\sqrt{R^2-a^2}-R^2\left(\frac{\pi}{2}-\arcsin\sqrt{1-(a/R)^2}\right)\right]\right\},$$

if the coordinate x is measured from the center of the drop. The local surface tension on the sphere $2\sigma_r$ leads to the following expression for the work of expansion of the drop:

$$W_4=\int_a^R 2\sigma_r\cdot 4\pi r^2\,dr=4\pi\sigma(R^2-a^2).$$

As a result, the energy balance under discussion can be written to within accuracy to the axial strain of the drop in the form $W_1+W_2+W_3=W_4$. Eliminating $b=l/R$ from the latter using (1), we obtain the following equation for determining η:

$$f(\eta,\varepsilon)=0,\quad q=1+(1+1/\varepsilon)^{-1},$$

$$\eta=a/R,\quad \varepsilon=\Delta l/2l,$$

$$f(\eta,\varepsilon)=\left(3-\frac{\pi}{4}\right)\eta^3+\eta^2+\frac{\eta}{2}$$

$$\times\left(\arcsin\sqrt{1-\eta^2}-\eta\sqrt{1-\eta^2}-4-\frac{\pi}{2}\right)+\frac{2}{3}q. \quad (2)$$

Since the experimental observations[2] attest to relative elongation of the contracted segment of the jet ($\Delta l>0$), the parameter ε is allowed to take only the positive values $\varepsilon>0$ ($q>1$), i.e., $\varepsilon=0$ is a lower bound for the possible location of the $\xi=f(\eta,\varepsilon)$ curve on the (η,ξ) plane. In the latter case, it takes the minimum ($f_\eta=0$, $f_{\eta\eta}>0$) value $\xi_m=f(\eta_*,0)\cong-0.05$ when $\eta=\eta_*\cong0.53$, so that the condition $\xi>\xi_m$ remains valid for any $\varepsilon>0$. Because of the numerical smallness of the value of ξ_m in comparison to the values of $\xi>2/3$ at the ends of the applicable range $0<\eta<1$ for the variation of η, Eq. (2) has two close roots, which merge at $\eta=\eta_*$, where the minimum on the $\xi=f(\eta,\varepsilon)$ curve in the limiting case of $\varepsilon=\varepsilon_*$ lies directly on the $\xi=0$ axis, and simultaneous satisfaction of the equality $f=f_\eta=0$ specifies the values $\eta=\eta_*\cong0.53$ and $\varepsilon_*\cong0.081$ ($q=q_*\cong1.075$).

If we take into account the formation of the tiny intermediate droplets upon rupture of the jet constriction, their volume should be added to the right-hand side of the mass balance equation (1). The latter evokes a certain increase in $q(\varepsilon)$ and, accordingly, a general upward displacement of the $\xi=f(\eta,\varepsilon)$ curves from the lower boundary $\xi_m<0$. Therefore, the narrow vicinity $\varepsilon\leqslant\varepsilon_*$ near the limiting value of this parameter $\varepsilon=\varepsilon_*$ ($q=q_*$), which can serve as an estimate of the mean diameter of the drops formed, will be predominantly realized. The resultant value of their diameter $2R=2a/\eta_*\cong1.89\cdot2a$ is in good agreement with the known results of the experimental observations for thin jets of low-viscosity liquids.[2] The use of Eq. (1) and the values of η_* and ε_* found gives $b=b_*\cong2.167$, which permits an estimate of the mean distance between the main drops of a jet which is breaking up:

$$L=2l+\Delta l=\frac{b_*}{\eta_*}(1+\varepsilon_*)\cdot2a\cong4.42\cdot2a.$$

This estimate also corresponds to the experimental data and is close to the wavelength λ_* given above.

[1] Lord Rayleigh, Proc. London Math. Soc. **10**, 4 (1879).
[2] L. Prandtl, *Essentials of Fluid Dynamics*, translated from the German [Blackie, London (1952); IL, Moscow (1951), 520 pp.].

Translated by P. Shelnitz

Formation of a high-current discharge plasma in metal vapor

M. A. Krasnogolovets

Institute of High-Energy Physics and Nuclear Physics, National Science Center, Kharkov Physicotechnical Institute, 310108 Kharkov, Ukraine
(Submitted November 10, 1998)
Zh. Tekh. Fiz. **69**, 134–137 (November 1999)

A method of converting the energy of a capacitive store into the energy of a high-current gas discharge in aluminum vapor is investigated. The spatial and temporal characteristics of the gas-discharge plasma are studied. The electron temperature, density of charged particles, and conductivity of the plasma are determined. © *1999 American Institute of Physics.* [S1063-7842(99)02611-2]

INTRODUCTION

High-temperature plasmas are used extensively in technological processes associated with modification of the composition and surface properties of materials, melting, hardening, and welding. The design of new switching devices, plasma amplifiers and oscillators, and high-intensity light sources would be unthinkable without the application of a high-temperature plasma as the active medium for coherent radiation. In this paper we give the results of experimental investigations of certain parameters of a plasma generated by the development of a high-current gas discharge in aluminum vapor.

The experiments were carried out on an apparatus designed for studying the processes of a recombining plasma in metal vapor. The main components of the apparatus are a discharge chamber, vacuum station, a device for controlling the discharge ignition time, and diagnostic equipment.

The stainless steel vacuum chamber has a volume of approximately $0.2\ \mathrm{m}^3$ and is evacuated to a residual gas pressure $p=10^{-5}$ Torr. Metal electrodes lead into the chamber through special insulating flanges to form the discharge gap. The length of the discharge gap can be varied from 1 cm to 10 cm. The electrodes are connected by a coaxial cable to an energy store in the form of a bank of UK 100-0.44H low-induction, high-voltage capacitors with a total capacitance $C=8\times10^{-6}$ F. The capacitor bank is charged by a high-voltage source, $U_0=(5-50)\times10^3$ V. The aluminum electrodes are the main providers of matter injected into the plasma. One electrode has a very simple structure and serves as the anode in the first half period of the discharge current. It comprises a cylinder of diameter $d=2\times10^{-2}$ m with a hemispherical end surface. The second, ignitor electrode (cathode) has the same outward appearance, but a complex internal structure. A plasma gun is installed inside the electrode for the ignition of electrical discharge between the main electrodes. When a short, high-voltage pulse of amplitude $u=2\times10^4$ V and duration $\tau=(2-3)\times10^{-6}$ s is fed to the ignitor, a discharge is triggered between the electrodes, initiating at the dielectric disk. The discharge runs along the barrel of the plasma gun and at the end of its path fires a slug of ionized gas into the discharge chamber, lowering the electric strength of the discharge gap and thus setting the stage for the development of a pulse arc discharge between the electrodes. The energy stored by the capacitor bank is injected into the plasma, heating it to high temperatures. Upon recombining, the plasma radiates over a wide range of wavelengths. Operation of the apparatus over a long period of time has demonstrated reliable operation of the ignition system.

The operation of the apparatus is synchronized with the diagnostic equipment by means of a G-5-15 pulse generator. The generator triggers the driven sweep of an S1-16 dual-beam oscilloscope and an S1-41 storage oscilloscope, and it generates a delayed pulse of positive polarity with an amplitude of 80–100 V and a duration of $(4-5)\times10^{-6}$ V, which inputs the firing time to the control unit.

A coaxial shunt or Rogowski loop is used to measure the current in the discharge circuit. Oscillograms of the voltage applied to the discharge electrodes are obtained by means of a resistor-capacitor voltage divider. The temporal characteristics of the plasma radiation in the x-ray region of the spectrum are investigated by means of an x-ray sensor consisting of: a cutoff filter (beryllium foil of thickness $d=30\times10^{-6}$ m), which does not transmit visible and ultraviolet radiation; a scintillator (NaI crystal) for the conversion of x-rays into visible light; and an FÉU-19M photomultiplier for the conversion of visible light into an electrical signal. The electrical signals from the coaxial shunt or Rogowski loop, the voltage divider, and the x-ray sensor are sent to the oscilloscope input.

EXPERIMENTAL INVESTIGATIONS OF THE PARAMETERS OF A HIGH-CURRENT GAS DISCHARGE

Comparatively accurate data on the electron temperature T_e can be obtained by a simple procedure based on measurement of the relative intensity of the undecomposed plasma x-rays transmitted through absorbing films of various thicknesses.[1]

If the wavelength dependence of the absorption coefficient μ for the selected filter is known, the relative attenuation of the intensity of the undecomposed plasma radiation spectrum is expressed by the equation[2]

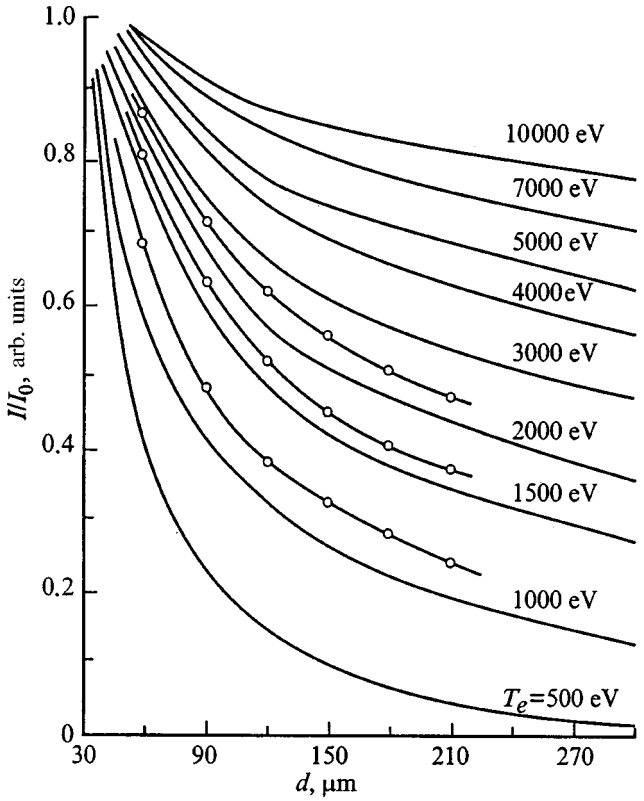

FIG. 1. Absorption curves for beryllium at various plasma electron temperatures, calculated (—) and experimental (–○–).

$$F_k(P_k)=\frac{2}{3}\int_{y_k}^{y_{k+1}}\left(1+\frac{1}{2}y\right)\exp\left(-\frac{P_k^4}{y^3}-y\right)dy,$$

$$y_k=\frac{h\nu_k}{T_e},\quad P_k=\left(\frac{a_k d h^3}{Te}\right)^{1/4},\quad a_k=\mu_k\nu^3,$$

where μ is the linear attenuation coefficient of the filter, d is the thickness of the absorbing filter, h is Planck's constant, and k and $k+1$ denote the absorption jumps in the vicinity of the resonance frequencies corresponding to atomic energy levels ($\nu_1<\nu_2<\ldots<\nu_n$).

We have used this equation to calculate the relative absorption coefficients of the beryllium foils at various temperatures T_e; the results are shown in Fig. 1. The experimental curves were obtained from measurements of the intensity of the radiation transmitted through a set of absorbing beryllium foils of thickness $30-210$ μm. The intensity of the transmitted radiation was recorded on RF-7 x-ray film, whose spectral sensitivity was essentially uniform over the investigated wavelength range. The results of photometry of the films have enabled us to determine the relative absorption coefficients of the beryllium foils at various temperatures T_e, and we have used these data to plot the experimental curves for three cases (at charging voltages $U_0=2\times10^4$ V, 3×10^4 V, and 4×10^4 V). Comparing the calculated and experimental curves, we find the electron temperature of the plasma for these cases: $u_0=2\times10^4$ V, 3×10^4 V, 4×10^4 V; $T_e\approx1200$ eV, 1600 eV, 2200 eV. It can be concluded from the results that a high-temperature plasma is formed by the development of a gas discharge in aluminum vapor. The electron temperature is used in determining the concentration of the plasma, its degree and multiplicity of ionization, and other parameters.

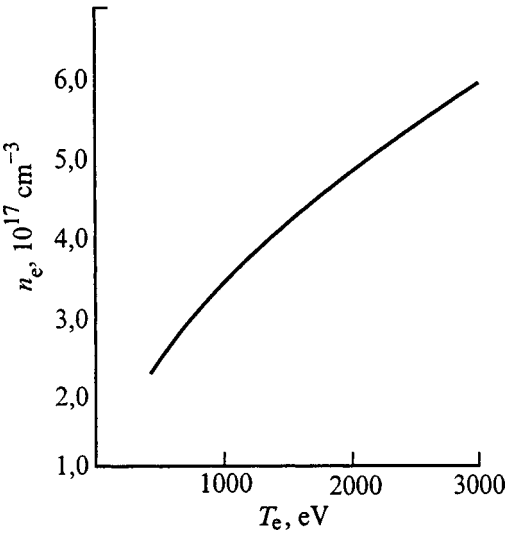

FIG. 2. Dependence of the electron density on the electron temperature.

No less important a parameter than T_e is the density of electrons n_e in the high-temperature plasma. We determine the density of electrons from the conditions of local thermodynamic equilibrium. In the local thermodynamic equilibrium model the distribution of electrons among the energy level is assumed to be governed entirely by collisions between particles, the collision processes occurring so frequently that after any change in the conditions in the plasma the appropriate distribution is established instantaneously. We base our calculations on the expression[3] $n_e\geqslant1.6\times10^{12}T_e^{1/2}x(p;q)^3$, where T_e is the electron temperature in K, and $x(p;q)$ is the potential energy of excitation from the qth level to the pth level in eV. For aluminum we have $x(p;q)=4.0$ eV.

The results of the calculations are shown in Fig. 2. They show that a plasma with an electron density $n_e\geqslant10^{17}$ cm^{-3} is formed during the development of a high-current gas discharge in aluminum vapor.

The conductivity of the plasma can be determined from the Spitzer equation[4]:

$$\sigma=\frac{3.3\times10^2\gamma(Z)T_e^{3/2}}{z^2\ln\Lambda},\quad \ln\Lambda=\ln\frac{3(kT_e)^{3/2}}{2(4\pi)^{1/2}z^3e^2n_e^{1/2}},$$

where Z is the average charge of the plasma ions, $\gamma(Z=1)=0.58$, $\gamma(Z=2)=0.68$, $\gamma(Z=3)=0.78$, Λ is the Coulomb logarithm, and e is the electron charge.

At high temperatures we have essentially 100% ionization of the plasma with an ionization multiplicity $Z=2$. The results of the calculations are shown in Fig. 3, from which it is evident that the discharge in aluminum vapor generates a plasma with a conductivity $\sigma>10^{11}$ S/m. This result reflects the very promising outlook of the plasma for energy transport applications.

The time behavior of the plasma has been investigated in the visible x-ray range of the radiation spectrum. The x-rays

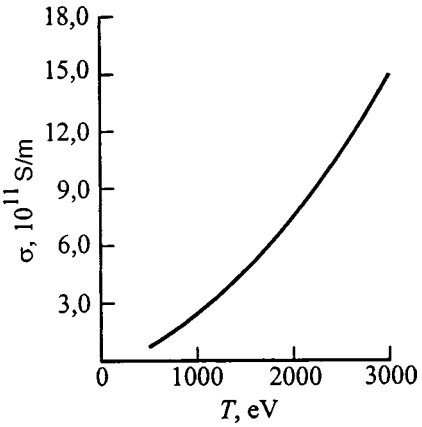

FIG. 3. Dependence of the conductivity of the plasma filament on the electron temperature.

were converted into visible light by means of a crystal scintillator and were recorded by a photomultiplier. Figure 4 shows an x-ray oscillogram for $U_0=35\times10^3$ V and a residual gas pressure in the chamber $p=10^{-5}$ Torr.

The jagged appearance of the oscillogram can be attributed to the onset of instability in the discharge,[5] which induces current breaks and an x-ray burst typical of this situation. The x-ray dose rate was measured by means of dosimeters from a KID-2 kit and was found to be fairly high. For example, at a discharge voltage of no more than 2.5×10^4 V on each side of the capacitor bank the dose rate is $P=2800$ R/s. Knowing the x-ray dose rate, we can determine the x-ray intensity $I=0.11P/\gamma$, where the numerical coefficient 0.11 is the roentgen energy equivalent in air, P is the x-ray dose rate, and $\gamma=6\times10^2$ 1/cm^{-1}. The x-ray intensity is therefore equal to $I=0.52$ W/cm^2.

We have also investigated the intensity of x-rays from the residual gases in the discharge chamber. We find that the x-ray intensity decreases as the residual gas pressure increases, and at a pressure $p_0=2\times10^{-2}$ Torr x-rays do not occur. The duration of radiation in the visible part of the spectrum is also equal to $(50-60)\times10^{-6}$ s. The time dependence of the radiation intensity is similar in its behavior to the dependence of the radiation intensity in the x-ray regions of the spectrum.

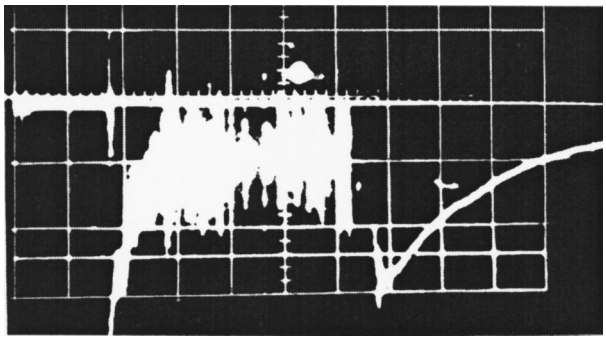

FIG. 4.

FIG. 5.

The spatial characteristics of the discharge have been investigated by means of a camera obscura in the visible and x-ray regions of the spectrum. An analysis of photographs of vacuum discharges shows that numerous cathode spots are observed on the electrodes. An intensely radiating plasma cloud is formed near the anode.

Figure 5 shows a discharge in x-rays, which we were able to use to determine the volume of the hottest plasma, which was located near the anode. This volume is approximately 4 cm^3 and radiated intensely in the x-ray wavelength range.

[1] S. Yu. Luk'yanov, *Hot Plasma and Controlled Nuclear Fusion* [in Russian], Nauka, Moscow (1975), 355 pp.
[2] V. F. Aleksin and V. A. Sutrunenko, Zh. Tekh. Fiz. **35**, 1945 (1965) [Sov. Phys. Tech. Phys. **10**, 1498 (1965)].
[3] R. H. Huddlestone and S. L. Leonard (Eds.), *Plasma Diagnostic Techniques* [Academic Press, New York, 1965; Mir, Moscow, 1967, p. 515].
[4] V. D. Zvorykin, A. D. Klementov, N. G. Kulinovskiĭ, and V. B. Rozanov, Kvantovaya Élektron. **3**, 131 (1976) [Sov. J. Quantum Electron. **6**, 182 (1976)].
[5] Yu. G. Gusev, *Handbook of Radiation Safety* [in Russian], Atomizdat, Moscow (1968), 286 pp.

Translated by James S. Wood

Synthesis of epitaxial layers of zinc oxide on nonorienting substrates

B. M. Ataev, I. K. Kamilov, A. M. Bagamadova, V. V. Mamedov, A. K. Omaev, and M. Kh. Rabadanov

Institute of Physics, Dagestan Scientific Center, Russian Academy of Sciences, 367003 Makhachkala, Russia
(Submitted January 11, 1999)
Zh. Tekh. Fiz. **69**, 138–140 (November 1999)

The first experiments on the growth of single-crystal layers of zinc oxide on nonorienting substrates (crystalline leucosapphire and fused quartz) by chemical transport reactions in a reduced-pressure flow-through reactor in a hydrogen atmosphere is reported. To ensure autoepitaxy on a surface of a nonorienting substrate, an optimized intermediate layer of zinc oxide of thickness $200-1000\,\text{Å}$, which provides a texture of basal orientation regardless of the orienting properties of the substrate, is preliminarily deposited by magnetron sputtering. It is shown that the subsequent growth of layers on such a surface by a chemical transport reaction to a thickness of 1–5 ensures high structural perfection, uniformity, and a very smooth surface, while polycrystalline films are deposited on the portion of the surface without a buffer layer. The proposed method can be used to grow heteroepitaxial structures and other electronic materials on nonorienting substrates using chemical transport reactions. © *1999 American Institute of Physics.* [S1063-7842(99)02711-7]

The synthesis of epitaxial layers on nonorienting substrates, i.e., surfaces of ceramics, glass, fused quartz, refractory metals, etc., or on surfaces of semiconductors with highly different lattice constants is a fairly complex, but attractive problem for many applied areas of electronics.

Approaches using graphoepitaxiy (artificial epitaxy),[1] particularly the synthesis of zinc oxide epitaxial layers on the amorphous surface of silicon wafers with a relief in the form of a one-dimensional lattice,[2] are known. The use of thin intermediate layers obtained by magnetron sputtering (the method produces textured layers of basal orientation with good adhesion even on nonorienting substrates) for improving the structure and uniformity of zinc oxide epitaxial layers on sapphire was reported in Ref. 3, where it was shown that a significant effect can be achieved even in the case of oriented substrates due to activation of an autoepitaxy mechanism. We do not know of any other work in which epitaxial layers were successfully synthesized on nonorienting substrates.

In the present work we used thin $(200-1000\,\text{Å})$ textured ZnO intermediate layers obtained by magnetron sputtering for the purpose of growing high-quality ZnO epitaxial films on nonorienting substrates (crystalline leucosapphire and fused quartz) in a reduced-pressure flow-through reactor by a chemical transport reaction. There was also additional interest in studying the possibility of stabilizing the growth of layers with the basal orientation in the case of chemical transport, since the synthesis of zinc oxide epitaxial layers of this orientation even on such well-oriented substrates as (0001) Al_2O_3 meets with certains difficulties.[4]

For the experiment we used $20\times15\,\text{mm}$ Polikor (a crystalline leucosapphire) and fused quartz substrates as materials which can withstand temperatures above 900 K and have high mechanical properties. The intermediate layers were deposited by dc magnetron sputtering using tablets, i.e., targets with a diameter of 4 cm and a thickness of 2–3 mm. For the purpose of ensuring stable discharging during dc magnetron sputtering in an atmosphere with $Ar:O_2=4:1$, we added 1 wt. % Ga_2O_3 to the high-purity ZnO power (this is not required for rf sputtering), the mixture was thoroughly stirred, and after being pressed, the targets were annealed at 1400 K for 10–12 h. The discharge current did not exceed 100 mA, and the thickness of the layers was determined from the sputtering time and amounted to 200–1000 Å. The intermediate layer was deposited on only half of the area of the substrate. This facilitated comparative measurements and eliminated the inaccuracies associated with the possible irreproduciblity of the technological parameters in the case of separate growth of the layers using a buffer layer and without one. After the intermediate layer was deposited, the substrates were transferred to the reduced-pressure flow-through reactor described in Ref. 5, where a method for optimizing the temperature regimes in the vaporization and deposition zones was given. Tablets of high-purity ZnO with a diameter of 3 cm and a length of 3 cm were used. The thickness of the epitaxial layers reached $1-5\,\mu\text{m}$. The structural perfection was monitored by x-ray diffraction and electron diffraction methods. The features of the exciton luminescence from both portions of the substrate, i.e., the portion with a polycrystalline layer and the portion with an epitaxial layer on the optimized buffer layer, were also studied.

The x-ray diffraction experiments were performed on a DRON-2 diffractometer using Cu K_α radiation, which was monochromatized using a pyrographite crystal. Figure 1 presents the diffraction patterns of the zinc oxide layers. The diffraction pattern in Fig. 1a was obtained from the portion of the substrate with a preliminarily deposited buffer layer, and the diffraction pattern in Fig. 1b was obtained from the

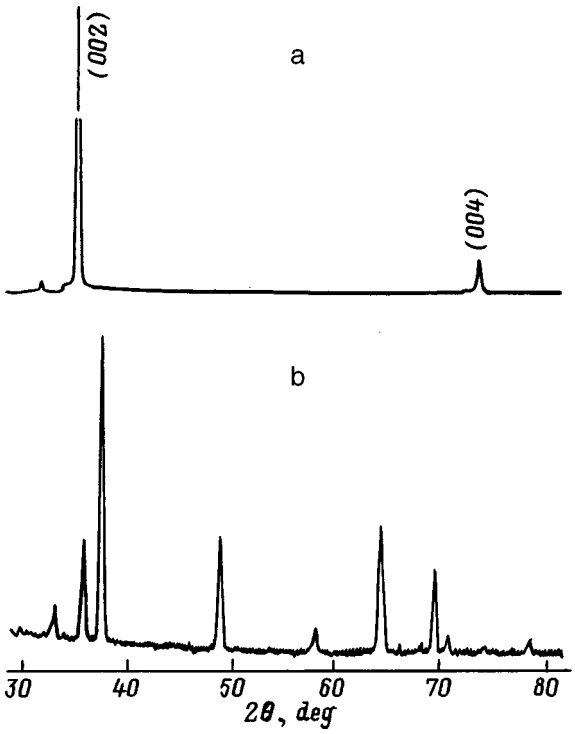

FIG. 1. Diffraction patterns of ZnO films obtained using a chemical transport reaction on a buffer layer (a) and on a clean fused quartz surface (b).

portion of the substrate without a buffer layer on the clean fused quartz surface. It can be seen that diffraction pattern *a* corresponds to highly oriented layers with the basal orientation, while diffraction pattern *b* is characteristic of polycrystalline zinc oxide films. The presence of only a series of (00*l*) reflections on the diffraction pattern in Fig. 1a does not yet permit drawing a final conclusion regarding the single-crystal nature of the layers, since reflections appear only from blocks parallel to the film surface in the focusing method used. Therefore, the layers were additionally investigated on an ÉMR-100 electron-diffraction camera. A typical electron diffraction pattern is shown in Fig. 2. It was concluded on the basis of the combined investigations that the films obtained using buffer layers are single-crystal films. A study of the rocking curves of the (002) peak showed that the tilt angle of the blocks does not exceed 0.5°. We also

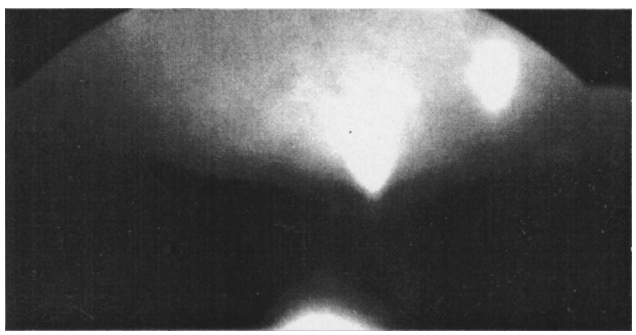

FIG. 2. Electron diffraction pattern of a ZnO epitaxial layer on a fused quartz substrate with a buffer layer.

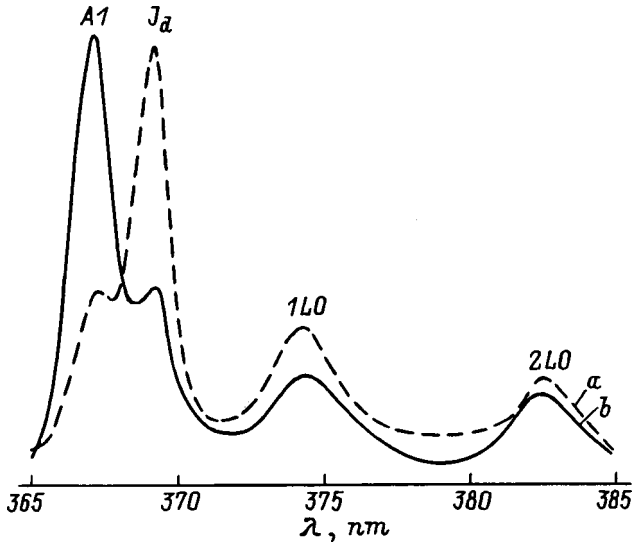

FIG. 3. Exciton photoluminescence spectra of ZnO films obtained on a buffer layer (a) and on a clean fused quartz surface (b).

note that the results of the structural analysis demonstrate the comparable quality of the layers on a nonorienting substrate with an intermediate coating even in comparison to an orienting substrate [see, for example, Ref. 6, which describes (1120) ZnO/(0112) Al_2O_3 layers obtained by magnetron sputtering]. An examination of the surface morphology corroborates the results obtained.

There is separate interest in an investigation of the UV photoluminescence of ZnO epitaxial layers, since the fine-structure features of the exciton spectra provide additional information on the perfection of the crystal lattice (see, for example, Ref. 7). The UV photoluminescence spectra recorded at 77 K from both portions of a fused quartz substrate, i.e., the epitaxial layer obtained using a buffer layer and the polycrystalline zinc oxide film, were studied. Figure 3 presents the corresponding spectra, which were obtained on a standard instrument using an SPM-2 monochromator. Spectrum *a* is characteristic of a perfect crystal structure and has predominant emission in the line for free exciton A1 and less intense emission in the band of bound exciton J_d (similar spectra were studied and identified in Ref. 7). Spectrum *b* corresponds to polycrystalline films and is distinguished by an intense band at 369 nm, while free-exciton emission is generally not observed. The 1LO/2LO intensity ratio and the half-widths of these lines also corroborate what has been said, i.e., spectrum *b* corresponds to a significantly more imperfect film.

Thus, not only does the use of intermediate layers permit the synthesis of highly oriented layers on nonorienting substrates, but also the quality of these layers is significantly higher than the quality of films obtained by magnetron sputtering. In our opinion, the slight complication of the growth procedure associated with the two-step process is compensated by the achievement of perfection which is not provided by chemical transport reactions and magnetron sputtering used separately. We also assume that more perfect zinc oxide

epitaxial layers can be obtained on nonorienting substrates after further optimization of the parameters of the buffer layers. We shall report the results of our work in this direction, as well as the use of buffer layers on other types of nonorienting substrates, in the future.

In conclusion, it should be noted that we do not see any reason why the proposed method could not be used to obtain epitaxial layers of other electronic materials on nonorienting substrates using chemical transport reactions.

This work was carried out with financial support from the Russian Fund for Fundamental Research (Grant No. 98-02-16141).

[1] N. N. Shevtal', *Rules Describing Real Crystallization Processes and Some Principles for Growing Single Crystals. Crystal Growth, Vol. 10* [in Russian], Nauka, Moscow (1974).
[2] V. I. Klykov and N. M. Gladkov, Izv. Akad. Nauk Latv. SSR, Ser. Fiz. Tekh. Nauk, No 1, 92 (1985).
[3] T. Shiosaki, S. Ohnishi, and A. Kawabata, J. Appl. Phys. **50**, 3113 (1979).
[4] S. A. Semiletov, A. M. Bagamadova, G. F. Kuznetsov *et al.*, Kristallografiya **23**, 377 (1978) [Sov. Phys. Crystallogr. **23**, 193 (1978)].
[5] A. Kh. Abduev, B. M. Ataev, and A. M. Bagamadova, Izv. Akad. Nauk SSSR, Neorg. Mater., No. 11, 1928 (1987).
[6] T. Shiosaki, S. Ohnishi, Y. Nurakami *et al.*, J. Cryst. Growth **45**, 346 (1978).
[7] A. Kh. Abduev, A. D. Adukov, B. M. Ataev *et al.*, Opt. Spektrosk. **50**, 1137 (1981) [Opt. Spectrosc. **50**, 626 (1981)].

Translated by P. Shelnitz

Current relaxation in microporous silicon

Yu. A. Vashpanov and V. A. Smyntyna

I. I. Mechnikov Odessa State University, 270026 Odessa, Ukraine

Khalmurat Azat

Xinjiang University, China

(Submitted January 11, 1999)
Zh. Tekh. Fiz. **69**, 141–142 (November 1999)

Current relaxation in samples of microporous silicon is observed at rates that depend on the applied voltage and the composition of the gaseous atmosphere. A possible physical mechanism for this phenomenon is discussed. © *1999 American Institute of Physics.* [S1063-7842(99)02811-1]

A new area in the physics and technology of semiconducting materials, nanoelectronics, is now being established.[1] Porous silicon, whose physical properties are under intense study, occupies a special place in research on nanostructures.[2] The creation of new semiconductor devices based on porous silicon necessitates research on the electronic and adsorption properties of this material and is a pressing problem in both its applied and theoretical aspects.

Samples of porous silicon were prepared by anode electrochemical etching of ⟨111⟩ silicon crystals (type KDB-10) in a water solution of hydrofluoric acid at a current density of 10 mA/cm^2. Porous silicon layers with a thickness of 7 μm with an average porosity of 35% were obtained by etching for 10 min. Gas-permeable slotted contacts of aluminum were deposited onto the porous silicon surface.[3] The effect of adsorbed gases on the in-plane electronic parameters of the material was studied at temperatures around 293 K. The gas concentration in the measurement chamber was varied using series GR03M gas generators. The carrier gas was high purity nitrogen, which has no effect on the electrical properties of porous silicon.[4]

When a voltage above a threshold $U_0 \approx 0.4$ V was applied to these porous samples after preliminary thermal annealing in vacuum at 523 K, a relaxation rise of the current was observed with the temporal characteristics for increasing voltage shown in Fig. 1 (curves *1–4*). Injecting air into the measurement chamber caused an increase in U_0 and a drop in the rate of rise of the current for the same voltages (Fig. 2, curves *1–4*). Of the gases in the air, oxygen and nitrogen are electrically neutral for porous silicon and only water vapor affects the electrical conductivity of silicon.[3]

An electromotive force has been observed[5] and attributed to the presence of imbedded microfields in the structural inhomogeneities of the material. The magnitude this emf changes with illumination and the adsorption of polar gases. An imbedded field develops during anode electrochemical processing of silicon,[6] which causes strong fluctuations in the potential along the surface of the material.[7] When a charge of $10^{17}e$ (e is the electronic charge) is localized near the minima of the potential energy, substantial electric fields of 10^5 V/cm are created.[8] One of the features of small, ultradisperse particles is a reduction in the ability of electrons to shield an electrostatic field, so that the radius of interaction between electrons is observed to increase.[9]

The Hill model[9] gives the current as a function of the applied voltage for ultradisperse media as

$$I(V,T) = \frac{1.24 \cdot 10^{11} c_p \sin h\left(\frac{eV}{kT}\right)}{(\Delta S)^2 m^*} \frac{\pi BkT}{\sin \pi BkT} \times \exp\left(-\frac{\delta E}{kT}\right) \exp(-1.03 \Delta S (m^* \varphi)^{1/2}),$$

where $\delta E = \delta E_0 - qF^{1/2}$ is the activation energy in a strong electric field, which depends on the magnitude of the electric field, and ΔS is the distance between the ultradisperse particles.

In our case, filamentary clusters of silicon are connected in series in the in-plane measurement regime. Because of the different sizes of the pores, the structure of porous silicon is extremely nonuniform along the surface. Size quantization of the electron energy spectrum makes the energy eigenvalues

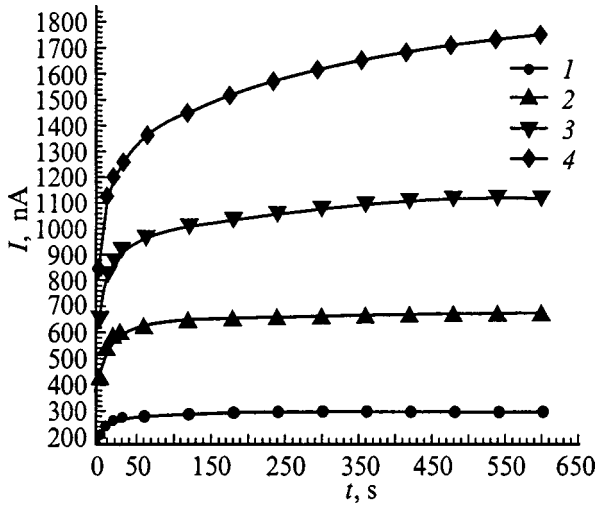

FIG. 1. Growth kinetics of the current in porous silicon structures in vacuum after annealing for voltages on the planar contacts of 0.6 (*1*), 1.2 (*2*), 1.8 (*3*), 2.4 V (*4*). The measurements were at room temperature.

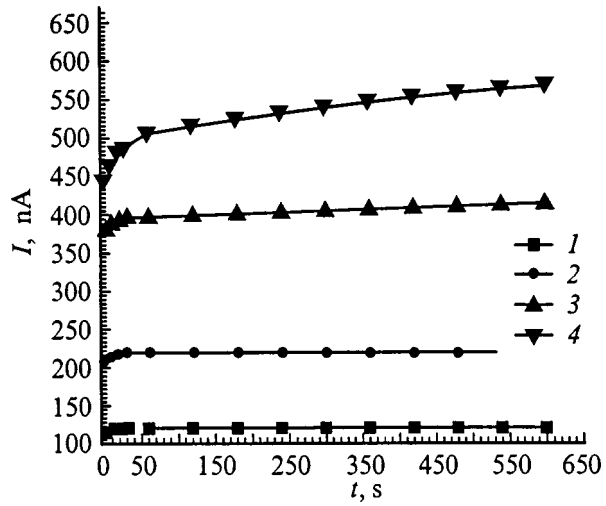

FIG. 2. As in Fig. 1 after air was let in.

of the electrons and holes depend on the transverse size of the filaments.[10] This leads to the appearance of local internal microbarriers which limit current flow.

The rise in the current when a voltage is applied to contacts on porous silicon could be explained by a drop in the parameter δE owing to a redistribution of the microfields in the material structure. When a polar gas is adsorbed, the structure of the microfields in microporous silicon changes.[4] In this case, $\delta E_0 - qF^{1/2}$ will vary to a small extent because of the increase in δE_0.

When a voltage appears at the contacts, because of the inhomogeneous structure of porous silicon the voltages falls on the individual microscopically inhomogeneous segments with large fluctuations in the potential, which limits current flow along the structure of the porous silicon. For inhomogeneities with dimensions on the order of 10 nm and $U = 0.1$ V, the electric field reaches 10^5 V/cm, which is comparable to the magnitude of the internal fields in porous silicon. In the regions where the voltages at the contacts, U, are less than U_g, there is essentially no change in δE since the electric field is not sufficient to change δE_0. In the region where the voltages contacts, U, exceed U_g, there is a drop in the barrier and the field is redistributed over the structure of the material. According to this model, an increase in the voltage at the contacts should lead to a rise in the relaxation rate, which is experimentally observed (Fig. 1, curves *1–4*).

The adsorption of water vapor changes the internal microstructure of the fields in the material.[5] Here the relaxation of the current for the same values of the voltage is much less (Fig. 2, curves *1–4*). This is possible because of the growth in δE_0.

In inhomogeneous porous silicon samples in-plane measurements reveal a kinetics for the rising dark current which is physically related to a readjustment of the internal microstructure of the fields in the material that develop during anode electrochemical processing of the silicon. The current relaxation process is affected by adsorption of polar water molecules, which create additional local microfields in the structure of the porous silicon.

[1] G. Abstreiter, *European Conference MRS'98*, Strasbourg, France (1998), PS-3.
[2] A. I. Yakimov, N. P. Stepin, and A. V. Dvurechenskiĭ, *European Conference MRS'98*, Strasbourg, France (1998), Vol. B11, p. 42.
[3] Yu. A. Vashpanov, Fotoélektronika, No. 6, pp. 68–71 (1996).
[4] Wu-Mian Shen, M. Tomkiewicz, and C. Lévy-Clément, J. Appl. Phys. **74**, 3635 (1994).
[5] Yu. A. Vashpanov, Pis'ma Zh. Tekh. Fiz. **23**(11), 77 (1997) [Tech. Phys. Lett. **23**(6), 448 (1997)].
[6] M. E. Kompan, I. Yu. Shabanov, Fiz. Tekh. Poluprovodn. **29**, 1859 (1995) [Semiconductors **29**, 971 (1995)].
[7] M. E. Kompan, I. I. Novak, and I. Yu. Shablanov, Fiz. Tverd. Tela (St. Petersburg) **37**, 678 (1995) [Phys. Solid State **37**, 195 (1995)].
[8] Yu. A. Vashpanov, *International Conference on Electron Localization and Quantum Transport in Solids*, Jaszowic, Poland (1996), pp. 177–178.
[9] I. D. Morokhov, L. N. Trusov, and V. N. Lapovok, *Physical Phenomena in Ultradisperse Media* [in Russian], Énergoizdat, Moscow (1984), 224 pp.
[10] M. S. Bresler and I. N. Yassievich, Fiz. Tekh. Poluprovodn. **27**, 871 (1993) [Semiconductors **27**, 475 (1993)].

Translated by D. H. McNeill