Technical Physics, Vol. 49, No. 9, 2004, pp. 1101–1103. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 1–3.

Two Decades Following the Discovery of Thermally Stable Elastic Properties of $La_3Ga_5SiO_{14}$ Crystal and Coining of the Term "Langasite" (A Review)

I. A. Andreev

Russian Research Center Vavilov State Optical Institute, St. Petersburg, 199034 Russia

Received March 9, 2004

Abstract—Twenty years ago, in December 2003, the Editorial Board of Technical Physics received a paper entitled "The New Piezoelectric Material "Langasite" $La_3Ga_5SiO_{14}$ with Zero Temperature Coefficient of the Elastic Vibration Frequency." This paper marked the beginning of a new stage in the development of piezo- and acoustoelectronics. During the twenty years following the publication of this paper, langasite crystals have led to the formation of a new trend in the preparation of materials with preset piezoelectric properties. Such an achievement in the field of piezoelectric materials science has been unheard of ever since 1921, when piezoelectric properties of α quartz were used to construct rf oscillators. Various piezoelectric devices, which form an inseparable part in the progress of modern communication devices, television, radars, and several kinds of defence technology are being produced by the thousands in developed countries all over the world. It is hard to overestimate the importance of producing miniaturized wideband filters based on new langasite-type materials for mobile communication to transmit video images in real time. This review contains a brief description of the initial stage of discovery and research of the unique piezoelectric properties of a langasite crystal and the fabrication of the first devices based on it. © *2004 MAIK "Nauka/Interperiodica".*

The advent of a new material in applied physics is a rare event. In piezoelectronics, dominated since 1921 until recently by the piezoelectric α-quartz crystal, which has unique mechanical and thermally stable properties, but is a weak piezoelectric, four new materials have appeared, viz., lithium tantalite $LiTaO_3$ (1966), berlinite $AlPO_4$ and fresnoite $Ba_2Si_2TiO_8$ (1976), and lithium tetraborate $Li_2B_4O_7$ (1983). These four crystals have two remarkable properties: (i) they have crystallographic orientations with a zero (at ~20°C) temperature coefficient of frequency (TCF) of elastic vibrations excited by a resonance electric field and (ii) the coefficient of electromechanical (piezoelectric) coupling, which determines the transmission bandwidth of a piezoelectric filter, is higher in these crystals than in α-quartz [1].

In the USSR, only lithium tantalite crystals were being developed in the 1980s to replace quartz in filters and tunable oscillators. The research was carried out at the Scientific–Industrial Complex (SIC) "Fonon"; no significant progress was made because of the technological problems caused by the ferroelectric properties of lithium tantalite crystals.

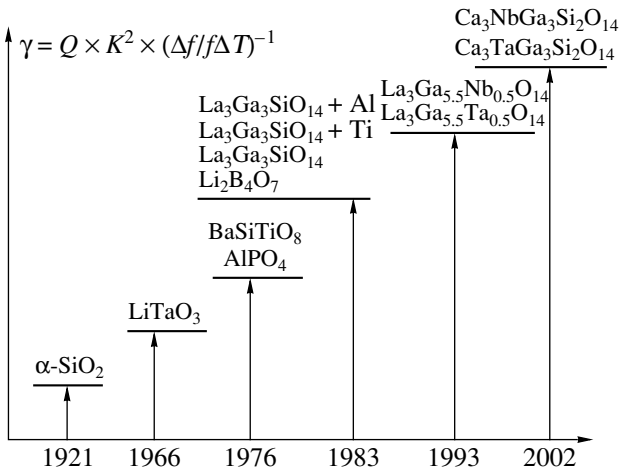
During the same period, leading Russian research centers started looking for new noncentrosymmetric crystals with a high optical nonlinearity to control the radiation frequency of solid-state lasers.

In 1982–1983, two groups of scientists in Kharkov (SIC "Monokristallreaktiv") and in Moscow (Moscow State University) were able to grow independently the first single crystals of lanthanum gallium silicate $La_3Ga_5SiO_{14}$ for this purpose. Unfortunately, the optical nonlinearity required for their practical application could not be detected in these crystals.

Apart from the optical nonlinearity studies of $La_3Ga_5SiO_{14}$ crystals, the elastic and piezoelectric properties of these materials were also studied at the Institute of Crystallography of the USSR Academy of Sciences and simultaneously at the Herzen State Pedagogical Institute, Leningrad. The discovery of piezoelectric properties in a material is not significant as such. It is well known that all centrosymmetric crystals possess these properties. The decisive factor for the practical application of piezoelectric crystal is the existence of crystallographic orientations along which the temperature coefficient of frequency of elastic vibrations is equal to zero at temperatures around 20°C. This is a consequence of the mutual compensation of positive and negative temperature coefficients of elastic moduli of the crystal and is observed very rarely in crystal physics.

The existence of crystallographic orientations with zero TCF in $La_3Ga_5SiO_{14}$ crystals was reported for the first time in [2]. It was shown experimentally for piezoelectric resonators with contour shear modes that *Y* cuts

1063-7842/04/4909-1101$26.00



Temporal dynamics of synthesis of materials with the best parameters for piezo- and acoustoelectronics.

of $La_3Ga_5SiO_{14}$ plates have parabolic temperature dependences of elastic vibration frequencies with temperature peaks T_0 in the interval 10-40°C, while the electromechanical parameters of $La_3Ga_5SiO_{14}$ are better than the analogous parameters for quartz. The lanthanum gallium silicate (LGS) crystal was termed langasite for the first time in this work. This term proved to be so appropriate that it has been used to denote these crystals all over the world ever since publication of the paper.

The first thermally stable orientations of LGS crystals were patented by Andreev and Dubovik, who received Inventors' Certificates for these materials on January 4, 1984, and March 4, 1984 [3, 4]. Since 1985, research and development of piezoelectric resonators based on these crystals and operating on bulk acoustic waves of various types have been carried out at the Special Design Office "Morion" (Leningrad). Inventors' Certificates were received [5–8] for optimal LGS orientations of resonators operating on shear, flexural, and longitudinal vibrations. A departmental commission of the Ministry of Electronic Industry, headed by V.B. Gruzinenko, leading specialist of the SIC "Fonon," acknowledged in February 1986 that "cuts with zero TCF have been obtained for flexural, longitudinal, and shear vibrations in lanthanum gallium silicate for the first time in the world…." Experimental samples of LGS resonators operating at various frequencies were sent to the Research Institute Dal'nyaya Svyaz' and the Television Research Institute, where they were tested successfully.

Simultaneously with the research and development of LGS resonators on bulk acoustic waves in the frequency range 50 kHz–20 MHz having a Q factor of up to 100 000, the first monolithic LGS filters for a frequency of 2.048 MHz (Leningrad Branch of the Communication Research Institute, 1985), the first delay lines operating on surface acoustic waves (SAWs) at a frequency of 88 MHz (SIC "Avangard," 1986) and SAW filters for a frequency of 36 MHz (Leningrad Institute of Aviation Instrumentation, 1986) were also developed at various research centers in Leningrad.

All these investigations showed that langasite crystals are promising basic materials for piezo- and acoustoelectronic devices (see, for example, [9]). It was also discovered that langasite doped with aluminium (langosital) or titanium (langositan) have the best piezoelectric and mechanical properties. These crystals were used for designing and fabricating resonators of longitudinal contour and thickness oscillations with zero TCF in the frequency range 130 kHz–1.8 MHz. The results of these investigations are protected by an Inventors' Certificate [10] dated August 10, 1987.

The research activity on LGS was halted at the SDO "Morion" in January 1987; however, the work in this field was intensified at SIC "Fonon."

At present, the research work is aimed at designing and fabricating piezoelectronic devices, resonators, filters, and delay lines, operating on bulk and surface elastic modes using langasite crystals and is being carried out intensely in Russia, France, Japan, the United States, Taiwan, etc. Apart from pure langasite and langasite doped with aluminum and titanium, two new materials (viz., langanite $La_3Ga_{5.5}Nb_{0.5}O_{14}$ and langatate $La_3Ga_{5.5}Ta_{0.5}O_{14}$), whose characteristics are better than or comparable to those of pure langasite have been developed. Both these materials are superior to piezoelectric α-quartz as regards piezoelectric communication, and their mechanical Q factor is comparable to that of α-quartz [11].

Four new derivatives of langasite have been synthesized quite recently: $Ca_3NbGa_3Si_2O_{14}$, $Ca_3TaGa_3Si_2O_{14}$, $Sr_3NbGa_3Si_2O_{14}$, and $Sr_3TaGa_3Si_2O_{14}$ [12]. The electromechanical coupling coefficient of these materials is four times larger and the acoustic losses are four times smaller than in α-quartz. Crystallographic orientations with zero TCF at room temperature have been discovered in $Ca_3NbGa_3Si_2O_{14}$ and $Ca_3TaGa_3Si_2O_{14}$ crystals.

It is expedient to present the dynamics of the formation of thermally stable piezoelectric single crystals for acousto- and piezoelectronics, where an arbitrary parameter γ is used as a quality criterion. This parameter is proportional to the mechanical quality factor Q, the square of the electromechanical coupling coefficient K, and is inversely proportional to the temperature coefficient of the frequency of elastic vibrations (see figure).

Thus, the parabolic dependence on the frequency of elastic vibrations with an extremum at room temperature in LGS plates discovered 20 years ago triggered a new trend in piezoelectronics, viz., the design,

research, and application of nonferroelectric langasite-type single crystals for rf stabilization and selection.[1]

ACKNOWLEDGMENTS

The author sincerely thanks G.I. Nadtochiĭ, a specialist at "Morion," for permission examine the materials of the Annual International Symposium on Frequency Control (USA) for the period 1994–2003.

REFERENCES

1. J. Zelenka, *Piezoelectric Sensors Using Resonators with Bulk and Surface Acoustic Waves* (Mir, Moscow, 1990) [translated from Czech].
2. I. A. Andreev and M. F. Dubovik, Pis'ma Zh. Tekh. Fiz. **10**, 487 (1984) [Sov. Tech. Phys. Lett. **10**, 205 (1984)].
3. I. A. Andreev and M. F. Dubovik, USSR Inventor's Certificate No. 1222170 (1985).
4. I. A. Andreev and M. F. Dubovik, USSR Inventor's Certificate No. 1230317 (1985).
5. I. A. Andreev, M. F. Dubovik, and V. L. Rassvetaev, USSR Inventor's Certificate No. 258101 (1987).
6. I. A. Andreev, M. F. Dubovik, and V. L. Rassvetaev, USSR Inventor's Certificate No. 258111 (1987).
7. I. A. Andreev and M. F. Dubovik, RF Patent No. 1373278 (1987).
8. I. A. Andreev and M. F. Dubovik, RF Patent No. 1382368 (1987).
9. M. F. Dubovik and I. A. Andreev, in *Proceedings of the IEEE Frequency Control Symposium, Boston, 1994.*
10. M. F. Dubovik, I. A. Andreev, *et al.*, USSR Inventor's Certificate No. 1506951 (1989).
11. R. G. Smythe *et al.*, in *Proceedings of the IEEE Frequency Control Symposium, Besancon, 1999*, pp. 816–820.
12. Shen Jen *et al.*, in *Proceedings of the IEEE Frequency Control Symposium, Kansas City, Missouri, 2000*, pp. 307–310.

Translated by N. Wadhwa

[1] This trend was formulated for the first time in the publication "Research and Development of New Strong Piezoelectrics and Prospective Piezoelectric Devices Based on Them with a Technical Level Higher than the Best Available Devices in the World," which was sent to a competition organized in September 1986 by SIC "Fonon" for the best pioneering research in the field of piezoelectronics.

Technical Physics, Vol. 49, No. 9, 2004, pp. 1104–1109. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 4–8.

THEORETICAL AND MATHEMATICAL PHYSICS

Scaling of 1/*f* Noise in Nonequilibrium Phase Transitions

V. P. Koverda and V. N. Skokov

Institute of Thermal Physics, Ural Division, Russian Academy of Sciences, Yekaterinburg, 620016 Russia

e-mail: vnskokov@itp.uran.ru

Received December 16, 2003

Abstract—Investigation of the dynamics of fluctuations of heat and mass transfer reveals that its crisis and transient modes exhibit high-energy pulsations with a power spectrum that is inversely proportional to frequency (flicker or 1/*f* fluctuations). Such a spectrum suggests energy transfer from high- to low-frequency modes and the possibility of large-scale catastrophic outbursts in the system being considered. The theory shows that such fluctuations arise in the system owing to the simultaneous occurrence of interacting phase transitions in the presence of white noise having a sufficiently high intensity. The distribution of fluctuations for scale transformations of the set of stochastic equations that describe the generation of 1/*f* noise is investigated. It is shown that, under a scale transformation, the Gaussian distribution of a random process having a 1/*f* spectrum passes to an exponential distribution, which is characteristic of the statistics of extreme outbursts. The probability of such outbursts must be taken into account in predicting the stability of various heat-transfer modes. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Heat and mass transfer in two-phase systems is characterized not only by the mean values of the parameters of the process being considered but also by chaotic fluctuation deviations from the mean values of these parameters. The enhancement of fluctuations is especially pronounced in critical and transient modes of heat and mass transfer [1]. The dynamics and evolution of random pulsations can be characterized by the frequency dependence of the spectrum of the fluctuation power. This dependence is determined by the ratio of the mean square of the noise-signal amplitude (and, hence, the noise-signal power) in the vicinity of the frequency f to the frequency bandwidth Δf. The spectrum where the power is bounded in the region of low frequencies corresponds to stable processes of heat and mass transfer. Experimental investigation of the dynamics of fluctuations of heat and mass transfer in crisis and transient modes [2–4] reveals that the low-frequency asymptotic behavior of the spectra in question is not always such. Low-frequency high-energy pulsations whose power spectra are inversely proportional to frequency (flicker or 1/*f* pulsations) are observed in crisis modes of boiling, in an explosive boiling of jets in overheated liquids, in vibrational modes of burning, and in an arc electric discharge. A feature peculiar to systems involving flicker noise (1/*f* noise) is that a considerable part of the energy of fluctuations is associated with very slow processes; moreover, formidable catastrophic outbursts are possible in such systems.

The power spectrum of fluctuations that is inversely proportional to frequency, $S \sim 1/f$, occurs in various physical, chemical, and biological systems [6, 7]. The 1/*f* behavior persists for fluctuations of power varying by a few orders of magnitude. By way of example, we indicate that 1/*f* pulsations of radiation from quasars and solar spots are known in astrophysics and that 1/*f* spectra are used to describe earthquakes and flooding in geophysics. In biology, one observes 1/*f* spectra in variations of insulin in the blood of diabetics and in cardiac and cerebral rhythms in the case of some diseases. In economics, financial flows and variations in the rates of exchange of shares obey the 1/*f* spectral dependence; also, flicker fluctuations occur in variations in the number of cars in traffic and even in music and speech [8].

In the literature, fluctuation processes characterized by a power spectrum that is inversely proportional to $1/f^{\alpha}$, where the exponent α varies within some range ($0.8 < \alpha < 2$), are most often classed with 1/*f* noise. A dynamical scaling that is observed at equilibrium critical points is a well-known property of $1/f^{\alpha}$ fluctuations. A great number of attempts have been made to explain a mechanism that could be responsible for the generation of scale-invariant fluctuations. The self-organized-criticality concept [9], which is used to describe complex systems featuring fully developed fluctuations, provides an outstanding example of this. In a state of self-organized criticality, the system being considered has a large number of metastable states in which it can be. In the process of its evolution, the system is self-organized and is tuned to a critical behavior involving scale-invariant fluctuations. In [9], self-orga-

1063-7842/04/4909-1104$26.00

nized criticality was demonstrated for the "sand pile" model of cellular automata. In the case of a constant flux of particles of sand in a sand pile that has reached self-organized criticality, there arise avalanches of different size or fluctuations, which maintain the critical state of the system irrespective of the magnitude of external effects. The theory of self-organized criticality leads to a spectrum of the form $1/f^{\alpha}$ ($\alpha \simeq 1.4$–2) and to a power-law distribution of fluctuations. An exact inverse proportionality to frequency ($\alpha = 1$) is observed for the spectrum of fluctuations of the voltage in the case where a current flows through a resistor [6, 10] and for nonequilibrium phase transitions interacting in heat- and mass-transfer processes [2–4].

In this article, we give a theory of fluctuation processes having a flicker power spectrum. According to this theory, flicker noise arises upon the superposition and interaction of nonequilibrium phase transitions. Along with the interaction, the fully developed fluctuation character of the process is a factor of importance here, manifesting itself in the generation of noise whose spectrum is divergent at low frequencies.

DISTRIBUTION OF 1/*f* FLUCTUATIONS

A theory of 1/*f* fluctuations in nonequilibrium phase transitions was proposed in [2]. The simplest stochastic equations describing the dynamics of fluctuations in a lumped system have the form

$$\begin{aligned} \frac{d\phi}{dt} &= -\phi\psi^2 + \psi + \Gamma_1(t), \\ \frac{d\psi}{dt} &= -\phi^2\psi + \lambda\phi + \Gamma_2(t), \end{aligned} \tag{1}$$

where ϕ and ψ are dynamical variables and Γ_1 and Γ_2 are Gaussian delta-correlated noises (white noise) characterized by identical variances. Owing to the coefficient $\lambda > 1$ of the variable ϕ in the second equation, the two equations in (1) are not equivalent. The appearance of this coefficient can be interpreted as the presence of macroscopic fluxes in the set of Eqs. (1). Below, we consider the case of $\lambda = 2$. Defining the potential as

$$\Phi = \frac{1}{2}\phi^2\psi^2 - \phi\psi, \tag{2}$$

we can recast the set of Eqs. (1) into the form

$$\begin{aligned} \frac{d\phi}{dt} &= -\frac{\partial\Phi}{\partial\phi} + \Gamma_1(t), \\ \frac{\partial\psi}{dt} &= -\frac{\partial\Phi}{\partial\psi} + \phi + \Gamma_2(t). \end{aligned} \tag{3}$$

In order to obtain deeper insight into the physical meaning of the potential (2), we make a linear transformation of our dynamical variables that corresponds to the rotation of the coordinates of the potential through an angle of $\pi/4$; that is, $\phi = \eta - \theta$ and $\psi = \eta + \theta$. In terms of the new variables, the potential assumes the form

$$\Phi = \frac{1}{2}\eta^4 - \eta^2 + \frac{1}{2}\theta^4 + \theta^2 - \eta^2\theta^2, \tag{4}$$

which is peculiar to the interaction of a subcritical and a supercritical phase transition (the order parameters are η and θ, respectively). The last term in Eq. (4) (it is proportional to $\eta^2\theta^2$) takes into account the interaction of the order parameters in the most general form.

In order to integrate the set of Eqs. (1) numerically, we rewrite it at $\lambda = 2$ as

$$\begin{aligned} \phi_{i+1} &= (\phi_i + \psi_i\Delta t)(1 + \psi_i^2\Delta t)^{-1} + \xi_i\Delta t^{0.5}, \\ \psi_{i+1} &= (\psi_i + 2\phi_i\Delta t)(1 + \phi_i^2\Delta t)^{-1} + \zeta_i\Delta t^{0.5}, \end{aligned} \tag{5}$$

where ξ_i and ζ_i are sequences of Gaussian random numbers having zero mean value and a standard deviation σ and simulating external white noise.

A feature peculiar to stochastic equations is that the time differential in the set of Eqs. (1) is of the second order of smallness in relation to the differential of a stochastic variable [11]. For this reason, the differentials $\xi_i\Delta t^{0.5}$ and $\zeta_i\Delta t^{0.5}$ in the set of Eqs. (5), which was written for the purposes of numerical integration, involve the time interval raised to a power of 0.5. This ensures mathematical conveniences in changing the integration step Δt: it is not necessary to correct the value of the standard deviation σ of the Gaussian random numbers ξ_i and ζ_i upon changing the integration step.

The set of Eqs. (1) and its computational version in (5) feature a noise-induced transition in the probability density $P(\sqrt{\phi^2\psi^2})$ [5]. The noise-induced transition means that, upon a change in the intensity of the external white noise, there occurs a change in the position of the extremum of the probability density (see Fig. 1). This change proceeds as a phase transition. If σ is less than some critical value σ_c ($0 < \sigma < \sigma_c$), the probability density $P(\sqrt{\phi^2\psi^2})$ has a maximum at some value of the argument $\sqrt{\phi^2\psi^2}$. At an external-noise intensity that corresponds to $\sigma = \sigma_c$, $P(\sqrt{\phi^2\psi^2})$ reaches a maximum at the origin of the coordinates. In response to a further increase in the intensity ($\sigma > \sigma_c$), the probability density behaves as a monotonically decreasing function (see Fig. 1).

At the external-white-noise intensity corresponding to the criticality of the noise-induced transition ($\sigma = \sigma_c$), the set of Eqs. (1) and, accordingly, the set of Eqs. (5) generate steady-state stochastic processes $\phi(t)$ and $\psi(t)$ whose power spectra exhibit the $1/f$ and the $1/f^2$ dependence, respectively. If the integration step is chosen in the interval $0.05 < \Delta t < 0.3$, the criticality of the noise-induced transition corresponds to the value of $\sigma_c = 0.8$; therefore, the spectral powers of the fluctua-

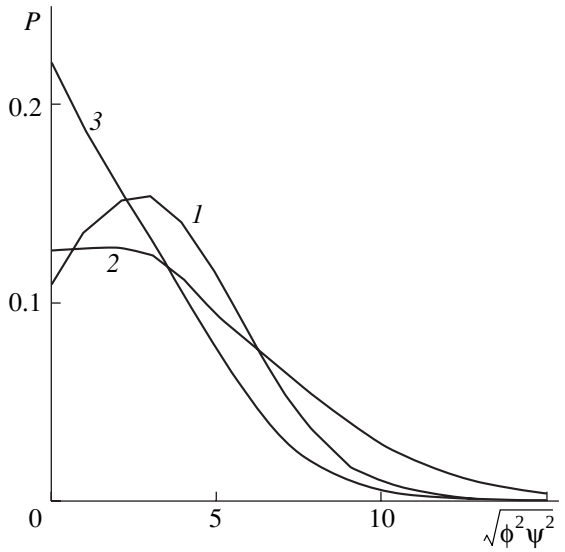


Fig. 1. Steady-state probability densities $P(\sqrt{\phi^2\psi^2})$ for the set of Eqs. (5) according to the calculation with an integration step of $\Delta t = 0.1$: (*1*) $\sigma < \sigma_c$, (*2*) $\sigma = \sigma_c$, and (*3*) $\sigma > \sigma_c$ ($\sigma_c = 0.8$).

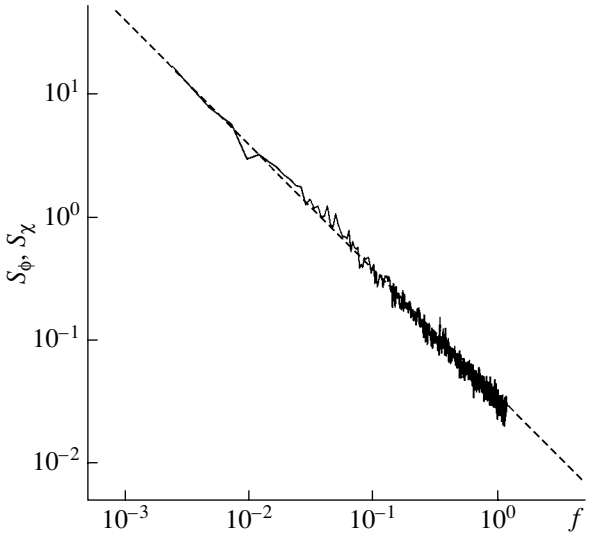


Fig. 2. Spectral density of the variables ϕ_i and χ_i. The dashed line represents a dependence of the form $S \sim 1/f$.

tions of ϕ_i and ψ_i follow the above dependences for intensities in the range $0.7 < \sigma < 0.9$.

The set of Eqs. (5) provides quite accurate results not only in the immediate vicinity of the critical point of the noise-induced transition but also over a rather broad range of the mean intensity of white noise [4]. This is due to a certain kind of consistency between the variables ϕ_i and ψ_i in the set of Eqs. (1). Even in [2], it was indicated for a lumped system that the mean value of the product of these variables is $\langle\phi\psi\rangle \simeq 1$, irrespective of initial conditions, the length of realization of a random process, and controlling parameters; being an invariant quantity that characterizes the system. Therefore, it is natural to expect that the inverse function of ψ also has a flicker power spectrum. In order to avoid a divergence at zero values of ψ_i, we specify the inverse function in the form [12] $\chi_i = \psi_i/(\varepsilon + \psi_i^2)$, where ε is a small parameter (usually, it is set to $\varepsilon \simeq 0.01$–0.02). The function χ_i is close to $1/\psi_i$ for the majority of the points of realization of a random process—only in the regions where ψ_i is close to zero is χ_i also close to zero. In specifying the inverse function for ψ_i, one can also use a different method to get rid of the divergence at the origin of coordinates, but this does not change the main result: the spectral density of the variable χ_i is in inverse proportion to the first power of frequency ($S_\chi \sim 1/f$) and is numerically consistent with the spectral density of the variable ϕ_i. The spectra of the variables χ_i and ϕ_i are given in Fig. 2. On the scale of the figure, they are indistinguishable and follow a $1/f$ dependence (dashed line in Fig. 2). In our calculations, we used 10^4 to 10^5 steps of integration of the set of Eq. (5) and performed averaging over a few tens of realizations. Although the power spectrum of the variable ψ_i is inversely proportional to the square of frequency, the power spectrum of the inverse variable χ_i is in inverse proportion to its first power. Thus, not only does the first equation in (1) or in (5) yield a $1/f$ spectrum, but the second equation generates such a spectrum upon passing from $\psi(t)$ to $\chi(t)$.

In contrast to the spectra, the distributions of the variables ϕ_i and χ_i are different. Figure 3 shows the distribution of the variable ϕ_i. This distribution is close to a Gaussian distribution, but it has long tails of large amplitude outbursts, these tails being noticeable only in semilogarithmic coordinates. Our numerical calculations revealed that the distribution of the variable ϕ_i can be approximated by the expression

$$P(\phi) = A\exp\left(-\frac{\phi^2}{2\sigma_\phi}\right) + B\exp\left(-\frac{|\phi|}{\sigma_\phi}\right), \qquad (6)$$

where A and B are constants ($A \gg B$) and $\sigma_\phi = 2\sigma dt^{0.5}$ is the standard deviation of the random process specified by $\phi(t)$. It is the second term in (6) that approximates the long tails against the background of the Gaussian distribution.

The distribution $P(\chi_i)$ differs from $P(\phi_i)$ by the presence of two maxima and a minimum at the origin (see Fig. 3), long tails being observed in this case just like for $P(\phi_i)$. However, they can be deformed here, depending on how one approximates the inverse function for ψ_i.

SCALING OF 1/*f* FLUCTUATIONS

We will now address the question of how the distributions of the variables change upon scaling realizations of random processes. For this purpose, we form,

from a sequence of calculated realizations $\{x_1, x_2, \ldots, x_N\}$, a sequence of roughened realizations $\{y^{(\tau)}\}$ by means of averaging over some time scale τ in accordance with the equation

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=\tau j}^{\tau(j+1)-1} x_i, \quad 0 \le j \le N/\tau, \tag{7}$$

where x_i is a stochastic variable (ϕ_i, χ_i, etc.).

The parameter τ is also referred to as the scale-transformation coefficient. For the first scale, the realization $\{y^{(1)}\}$ is merely an input realization. The length of each subsequent roughened realization decreases by a factor of τ; that is, it contains N/τ points. We note that this scale transformation does not change the frequency dependence of the spectrum. In the case of $\phi_i^{(\tau)}$ and $\chi_i^{(\tau)}$, the spectrum remains inversely proportional to frequency, $S \sim 1/f$.

The realizations of the input and roughened random processes are illustrated in Fig. 4. From this figure, one can see that, with increasing scale-transformation coefficient, the realizations of the random processes specified by $\chi_i^{(\tau)}$ and $\phi_i^{(\tau)}$ approach each other (see Fig. 4). At $\tau = 32$, the coefficient that characterizes the correlation of the realizations of the processes $\chi_i^{(\tau)}$ and $\phi_i^{(\tau)}$ is greater than 0.9. As the scale-transformation coefficient grows further, the realizations of $\chi_i^{(\tau)}$ and $\phi_i^{(\tau)}$ tend to each other.

The variations of the distributions of the variables $\chi_i^{(\tau)}$ and $\phi_i^{(\tau)}$ are shown in Fig. 5. At $\tau = 32$, the distributions for the two processes in question are nearly coincident and can be approximated by the formula

$$P(\chi) = C\chi^2 \exp\left(-\frac{|\chi|}{\sigma_\chi}\right), \tag{8}$$

where C is a normalization factor.

The exponential factor on the right-hand side of (8) describes long-wave outbursts of a random process having a $1/f$ power spectrum. This is in accord with the results of Antal *et al.* [13], who found that the scaling function for a roughened distribution of a certain class of periodic signals having a $1/f$ spectrum appears to be the distribution of extreme outbursts [14, 15]. The results of the present study reveal that, under a scale transformation, the Gaussian distribution of the stochastic process $\phi(t)$ having a $1/f$ spectrum passes to an exponential distribution, which is specific to the statistics of extreme outbursts. With increasing scale-transformation coefficient, the distribution of the variable χ_i tends to the same distribution at a much higher rate.

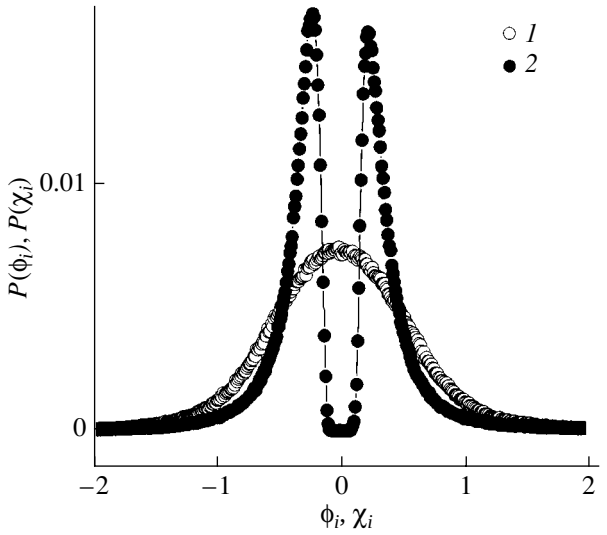


Fig. 3. Distributions of the variables being considered: (*1*) $P(\phi_i)$ and (*2*) $P(\chi_i)$.

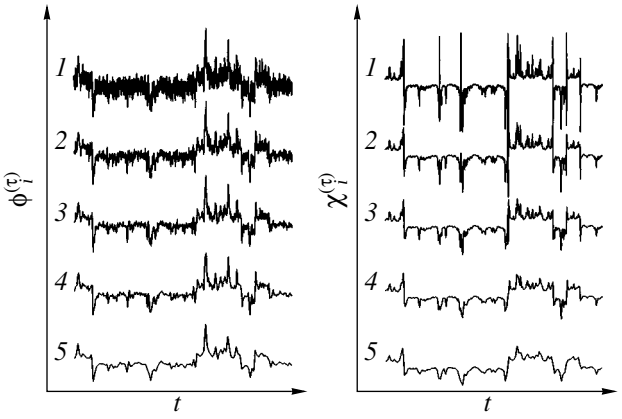


Fig. 4. Realizations of the input and roughened random processes $\phi_i^{(\tau)}$ and $\chi_i^{(\tau)}$ for the scale-transformation coefficient τ of (*1*) 1, (*2*) 4, (*3*) 8, (*4*) 16, and (*5*) 32.

ENTROPY ANALYSIS OF REALIZATIONS

In order to obtain a numerical characterization of the change in the distribution upon a scale transformation, we define the quantity

$$H(x) = -\sum_{x_i} p(x_i)\log(p(x_i)), \tag{9}$$

which has the meaning of information entropy. In [16], the use of the scale-factor dependence of the entropy was called a multiscaling entropy analysis of realizations of random processes. The results obtained by calculating the entropy of roughened realizations are given in Fig. 6 versus the scale-transformation coefficient τ. From this figure, one can see that, for the random process $\chi_i^{(\tau)}$, the entropy in question is independent of τ

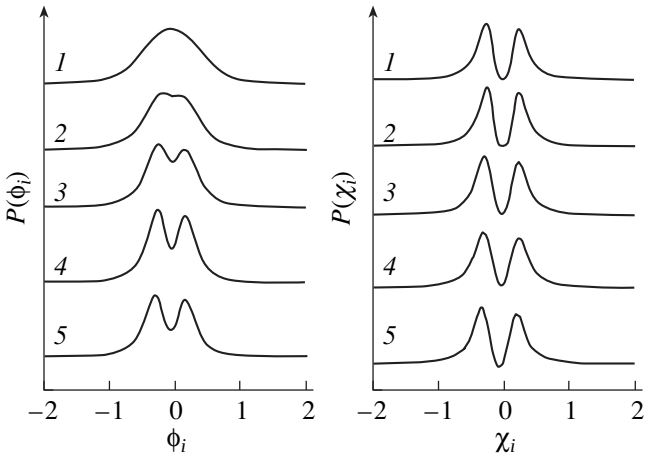


Fig. 5. Distributions of the variables $\chi_i^{(\tau)}$ and $\phi_i^{(\tau)}$ at various values of the scale-transformation coefficient τ: (*1*) 1, (*2*) 4, (*3*) 8, (*4*) 16, and (*5*) 32.

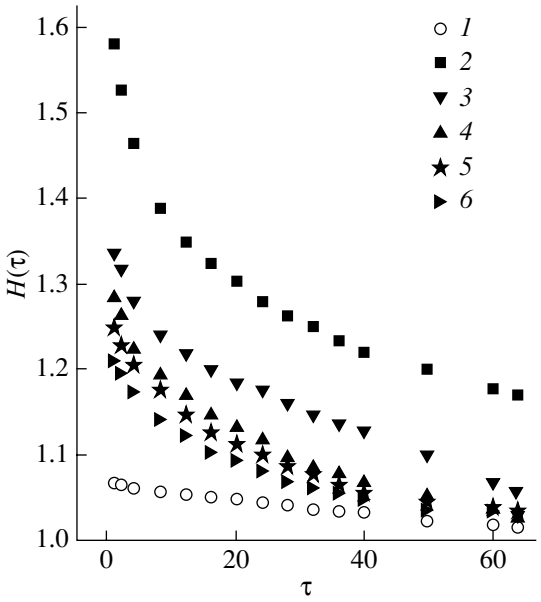


Fig. 6. Entropy $H(\tau)$ of roughened realizations for the random processes (*1*) $\chi_i^{(\tau)}$ and (*2–6*) $\psi_i^{(\tau)}$ as a function of the scale-transformation coefficient τ for various values of the controlling parameter λ: λ = (*2*) 4, (*3*) 3, (*4*) 2, (*5*) 1.8, and (*6*) 1.5.

(Fig. 6, points *1*), just like the distribution function itself, undergoing no changes. Nearly from the outset, the distribution function $P(\chi_i^{(\tau)})$ is scale-invariant, as in the process of self-organized criticality. As the scale-transformation coefficient increases, the random-process entropy $H(\chi_i^{(\tau)})$ decreases, tending to $H(\chi_i)$.

Under a scale transformation, the random-process entropy may characterize the effect of other controlling parameters of the system. In particular, an increase in the coefficient λ at the linear term on the right-hand side of the second equation in (1) in relation to (2) would lead to a prolongation of a Gaussian character of the process $\phi_i^{(\tau)}$ in response to the growth of the scale-transformation coefficient τ. For λ values smaller than that corresponding to (2), a scale-invariant distribution (close to the distribution of $\chi_i^{(\tau)}$) is obtained at smaller values of the scale-transformation coefficient. At $\lambda = 2$, the set of Eqs. (1) leads to $1/f$ fluctuations over the broadest region of white-noise intensities and integration steps. This is because of the maximal self-organization of random processes having a $1/f$ spectrum: a Gaussian process and a process characterized by a scale-invariant distribution function.

CONCLUSIONS

The theory of fluctuation processes having a flicker-type power spectrum in the presence of nonequilibrium phase transitions reveals that, at large values of the scale-transformation coefficient, the roughened distribution becomes exponential, which is peculiar to the statistics of extreme outbursts. The fully developed fluctuation character of heat- and mass-transfer processes in two-phase systems and the generation of noise whose spectra are divergent at low frequencies result in energy transfer from high- to low-frequency vibrations and may lead to large-scale outbursts in the system being considered. The possibility of such outbursts must be taken into account in assessing the stability of systems featuring fully developed fluctuation processes. Along with a spectral analysis, it is useful to perform a scale analysis of distribution functions.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 03-02-16215) and by the Program of Basic Research at the Russian Academy of Sciences (OEMMPU).

REFERENCES

1. S. S. Kutateladze and V. E. Nakoryakov, *Heat Exchange and Waves in Gas–Liquid Systems* (Nauka, Novosibirsk, 1984) [in Russian].
2. V. P. Koverda, V. N. Skokov, and V. P. Skripov, Zh. Éksp. Teor. Fiz. **113**, 1748 (1998) [JETP **86**, 953 (1998)].
3. V. N. Skokov, A. V. Reshetnikov, and V. P. Koverda, Teplofiz. Vys. Temp. **38**, 786 (2000).
4. V. N. Skokov, V. P. Koverda, A. V. Reshetnikov, *et al.*, Int. J. Heat Mass Transf. **46**, 1879 (2003).
5. V. P. Koverda and V. N. Skokov, Dokl. Akad. Nauk **386**, 187 (2002) [Dokl. Phys. **47**, 654 (2002)].
6. Sh. M. Kogan, Usp. Fiz. Nauk **145**, 285 (1985) [Sov. Phys. Usp. **28**, 170 (1985)].

7. M. B. Weisman, Rev. Mod. Phys. **60**, 537 (1988).
8. Yu. L. Klimontovich, *Statistical Theory of Open Systems* (Yanus, Moscow, 1995; Kluwer Academic, Dordrecht, 1995).
9. P. Bak, Ch. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
10. A. V. Yakimov and F. N. Hooge, Physica B **291**, 97 (2000).
11. A. I. Olemskoĭ, Usp. Fiz. Nauk **168**, 287 (1998) [Phys. Usp. **41**, 269 (1998)].
12. V. P. Koverda and V. N. Skokov, Dokl. Akad. Nauk **393**, 184 (2003) [Dokl. Phys. **48**, 615 (2003)].
13. T. Antal, M. Droz, G. Guorgui, and Z. Rach, Phys. Rev. Lett. **87**, 240601 (2001).
14. S. T. Bramwell, K. Christensen, J.-Y. Fortin, *et al.*, Phys. Rev. Lett. **84**, 3744 (2000).
15. K. Dahlstedt and H. J. Jensen, J. Phys. A **34**, 11193 (2001).
16. M. Costa, A. L. Goldberger, and C.-K. Peng, Phys. Rev. Lett. **89**, 068102 (2002).

Translated by A. Isaakyan

Technical Physics, Vol. 49, No. 9, 2004, pp. 1110–1115. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 9–14.

GASES AND LIQUIDS

Theoretical Analysis of the Formation of a Layered Ice Structure at the Surface of a Plate Placed in a Flow of a Supercooled Water Aerosol

R. G. Zakinyan

Stavropol' State University, Stavropol', 355035 Russia
e-mail: zakinyan@mail.ru

Received June 30, 2003; in final form, January 13, 2004

Abstract—The motion of the front of crystallization and the growth of a film at the surface of a plate are analyzed in the case of a laminar and in the case of a turbulent flow mode. Conditions are determined under which there occurs a transition from a matt inhomogeneous structure to a transparent homogeneous structure of ice. It is shown that, for a film to be steadily preserved at the plate surface, the film thickness must be larger than a critical equilibrium-thickness value $h_{b.c}$, in which case a transparent homogeneous structure of ice is formed. Otherwise, the film at the plate surface is unstable and disappears in the course of time. The icing of aircrafts is the most important application of the results obtained in this study. *© 2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is well known that, if a body is placed in a flow of a supercooled water aerosol, a liquid film is formed at the surface of the body, the character of crystallization there being dependent on the properties of the aerosol flow. For example, ice-cover structures of two basic types—a transparent homogeneous and a matt inhomogeneous structure—are observed in the icing of aircrafts in supercooled clouds [1].

A theory that describes the formation of a layered structure of ice growing in a supercooled-aerosol flow was first developed in [1] in considering the problem of the icing of aircrafts in supercooled clouds. According to this theory, a film is formed at the surface of a body placed in a flow of a supercooled water aerosol. Under the effect of the tangent stress generated by an air flow, the film begins to move. At realistic values of the airflow velocity and of the film thickness, the motion of the film can be either laminar or turbulent. This in turn leads to two mechanisms of heat transfer in the crystallization of the film, a molecular and a turbulent mechanism. In the former case (molecular mechanism), the film proves to be unstable—it disappears, with the result that drops undergo crystallization without merging, forming a matt inhomogeneous structure of ice. In the latter case (turbulent mechanism), crystallization proceeds under a film of steady-state thickness, with the result that there arises a transparent homogeneous structure of ice [2].

In [3–5], the mechanism proposed in [1] was considered under the assumption that the crystallization-front temperature T' is constant and is equal to the stable-crystallization temperature of T_0 = 273 K. This made it possible to obtain an analytic solution to the problem of the growth of ice under the film.

The icing of aircrafts is the most important application of the results obtained in the present study. In [6], the theory outlined here was formulated for the case of a spherical surface and was applied to explain the layered structure of a hailstone.

1. COEFFICIENT OF TURBULENT THERMAL DIFFUSIVITY

Before proceeding to calculate the rate of the growth of ice, we will preliminarily find an expression for the turbulent-thermal-diffusivity coefficient k_t. According to [7, 8], we have

$$k_t = b\frac{v^2 z^2}{\nu}, \quad v = \sqrt{\tau_0/\rho}, \tag{1}$$

where $b = 0.02$ is a dimensionless constant that is determined on the basis of experimental data, z is the distance (in meters) from the plate to an arbitrary level within the film, v is the dynamic velocity (in m/s units), ν is the water-molecular-viscosity coefficient (in m^2/s units), τ_0 is the friction tension at the plate surface (in N/m^2 units), and ρ is the water density (in kg/m^3 units).

Taking into account (1), we can find the turbulent-thermal-diffusivity coefficient k_t averaged over the entire film thickness $h = z'' - z'$. We have

$$\bar{k}_t = \frac{1}{z'' - z'}\int_{z'}^{z''} k_t dz = \frac{b}{2\nu} v^2 h^2, \tag{2}$$

1063-7842/04/4909-1110$26.00 © 2004 MAIK "Nauka/Interperiodica"

where z' and z'' are the distances (in meters) from the plate to, respectively, the crystallization front and the film surface, so that $h = z'' - z'$ is the film thickness, m.

The dynamic velocity v and the characteristic velocity u_∞ of the motion of the film are related by the equation

$$v = \sqrt{\frac{c_f}{2}}u_\infty, \quad (3)$$

where c_f is the drag coefficient, u_∞ being defined as the characteristic velocity (in m/s units) of the motion of the film at a large distance from the ice–water interface.

According to experimental data, the drag coefficient for a plate is given by [10]

$$\frac{c_f}{2} = 0.0294\mathrm{Re}^{-0.2}, \quad (4)$$

where $\mathrm{Re} = (u_\infty X)/\nu$ is the Reynolds number, X being the characteristic size (in meters) of the plate.

With allowance for relations (3) and (4), the expression for the dynamical velocity in the case of a plate takes the form [9]

$$v = 0.17\frac{u_\infty}{\mathrm{Re}^{0.1}}. \quad (5)$$

In order to find a relation between the velocity u_∞ of the flow of the film and the velocity of the air flow, we assume, according to Prandtl [11], that the turbulent friction τ_0 at the plate surface in the absence of the separation of drops from the film surface is equal to the friction F at the film surface; that is,

$$\tau_0 = 0.0294\mathrm{Re}^{-0.2}\rho u_\infty^2 = F = 0.0294\mathrm{Re}_a^{-0.2}\rho_a V^2, \quad (6)$$

where $\mathrm{Re}_a = VX/\nu_a$ is the Reynolds number for an air flow, ν_a is the kinematic-viscosity coefficient for air (in m^2/s units), ρ_a is the air density (in kg/m^3 units), and V is the air-flow velocity (in m/s units).

From (6), it follows that

$$\frac{u_\infty}{\mathrm{Re}^{0.1}} = \sqrt{\frac{\rho_a}{\rho}}\frac{V}{\mathrm{Re}^{0.1}}. \quad (7)$$

Substituting (7) into (5), we obtain

$$v = 5.2\times 10^{-3}\frac{V}{\mathrm{Re}_a^{0.1}}. \quad (8)$$

Substituting expression (8) for v into (2), we obtain the turbulent-thermal-diffusivity coefficient in the form

$$\bar{k}_t = B\mathrm{Re}^{-0.2}V^2h^2, \quad (9)$$

where $B = 0.15\ (m^2/s)^{-1}$ is a constant.

Having derived expression (9) for the coefficient of turbulent thermal diffusivity, we now proceed to calculate the rate of the growth of ice under the film.

2. GROWTH OF ICE UNDER THE FILM IN THE CASE OF A TURBULENT FLOW

The thickness of a film at the plate surface is determined by the difference of the velocities at which the crystallization front and the film surface move (dz'/dt and dz''/dt, respectively). In order to solve the problem at hand, it is necessary to impose two boundary conditions, that at the crystallization front and that at the film surface. In the case of a laminar mode of film motion, the equation of heat balance at the crystallization front z' has the form

$$\rho[L_c - c(T_0 - T_\infty)]\frac{dz'}{dt} = \lambda_i\left.\frac{\partial T}{\partial z}\right|_{z=z'} - \lambda\left.\frac{\partial T}{\partial z}\right|_{z=z'}, \quad (10)$$

where L_c is the specific heat of water crystallization (in J/(kg K) units); c is the specific heat capacity of water (in J/(kg K) units); λ_i and λ are the thermal conductivities (in J/(m s K) units) of, respectively, ice and water; $T_0 = 273$ K is the crystallization-front temperature; T_∞ is the ambient-medium temperature (in Kelvin degrees); and t is time (in seconds).

Equation (10) is the well-known Stefan condition [12].

Disregarding the removal of heat through the crystal, we can represent the heat-balance equation at the crystallization front

$$\rho[L_i - c(T_0 - T_\infty)]\frac{dz'}{dt} + \lambda\left.\frac{\partial T}{\partial z}\right|_{z=z'} = 0. \quad (11)$$

In a similar way, we can write the heat-balance equation at the crystallization front for the case of a turbulent mode of the film flow. Specifically, we have

$$\rho[L_i - c(T_0 - T_\infty)]\frac{dz'}{dt} + c\rho k_t\left.\frac{\partial T}{\partial z}\right|_{z=z'} = 0. \quad (12)$$

At the film surface, we impose the condition requiring that the heat flux to the film be equal to the heat removal from it (quasi-steady-state balance); that is,

$$-\lambda\left.\frac{\partial T}{\partial z}\right|_{z=z''} = \alpha(T_s - T_\infty) + L\beta(\rho_s - \rho_\infty), \quad (13)$$

where α is the coefficient of heat transfer between the film surface and ambient air (in $J/(m^2\ K)$ units); β is the coefficient of mass transfer between the film surface and ambient air (in m/s units); L is the specific heat of water evaporation (in J/kg units); and ρ_s and r_∞ are the steam densities (in kg/m^3 units) at the film surface and in the ambient medium, respectively.

The first term on the right-hand side of Eq. (13) represents the heat removed from the film surface via convective heat transfer, while the second term is the heat spent into evaporation from the film surface. For the turbulent mode of film motion, the coefficient λ in (13) must be replaced by $\lambda_t = c\rho k_t$. Under the condition in

(13), the temperature of the film surface remains constant, $T_s = T_1$.

We assume that a stationary temperature distribution is established in the film. We then have

$$\frac{dT}{dz} = \left.\frac{dT}{dz}\right|_{z=z'} = \left.\frac{dT}{dz}\right|_{z=z''} = -\frac{T_0 - T_1}{z'' - z'}. \quad (14)$$

The equation of mass balance at the film surface can be represented in the form

$$\frac{dz''}{dt} = \frac{qEV}{k_f \rho}, \quad (15)$$

where q is the water content in the flux (in kg/m^3 units), E is a coefficient that characterizes the capture of drops by the film surface (in the calculations, this capture coefficient was taken to be constant), and k_f is the shape coefficient ($k_f = 1$ for a plate).

In (15), the air-flow velocity is directed along the plate surface. The capture coefficient specifies the fraction of drops captured by the plate.

Taking into account relations (9), (12), (14), and (15), we represent the velocity of crystallization-front motion in the turbulent mode of film motion in the form

$$\frac{dz'}{dt} = \frac{ut - z'}{\tau}, \quad (16)$$

where

$$u = \frac{dz''}{dt}; \quad \tau = \frac{L_c c \rho (T_0 - T_\infty)}{cB\mathrm{Re}_a^{-0.2} V^2 \Delta T}; \quad \Delta T = T_0 - T_1. \quad (17)$$

Here, $u = dz''/dt$ has the meaning of the rate of the growth of the film owing to the influx of drops; at the values taken here for the velocity of the flow and its water content, u is constant.

In terms of the variable $h = ut - z'$, Eq. (16) assumes the form

$$\frac{dh}{dt} + \frac{h}{\tau} = u. \quad (18)$$

The steady-state film thickness h_e found from the condition $dh/dt = 0$ is given by

$$h_e = ut = \frac{qE[L_c - c\rho(T_0 - T_\infty)]}{c\rho B \Delta T \mathrm{Re}_a^{-0.2} V}. \quad (19)$$

Solving Eq. (18), we obtain

$$h = h_0 \exp\left(-\frac{t}{\tau}\right) + h_e\left(1 - \exp\left(-\frac{t}{\tau}\right)\right), \quad (20)$$

where h_0 is the initial film thickness (in meters).

With the aid of (18) and (19), we arrive at

$$\frac{dh}{dt} = u\left[1 - \frac{h}{h_e}\right]. \quad (21)$$

From (21), it follows that, if $h < h_e$, then $dh/dt > 0$, so that the film grows; but if $h > h_e$, then $dh/dt < 0$, with the result that the film diminishes. The same can be seen from the solution in (20). Irrespective of whether $h_0 > h_e$ or $h_0 < h_e$, the film thickness tends to h_e with time.

Let us introduce the notation $w = dz'/dt$. Solving Eq. (16), we obtain

$$w = w_0 \exp\left(-\frac{t}{\tau}\right) + u\left(1 - \exp\left(-\frac{t}{\tau}\right)\right), \quad (22)$$

where w_0 is the initial crystallization rate (in m/s units).

From (22), it can be seen that the velocity of crystallization-front motion tends to the velocity of film-surface motion with time, the film thickness remaining constant.

Thus, we see that, if the motion of the film is turbulent—recall that this mode is realized at Reynolds numbers above a critical value ($\mathrm{Re} \geq 1500$) [1, 2]—the growth of ice proceeds under a film of steady-state thickness, this leading to the formation of a transparent homogeneous structure of ice.

3. GROWTH OF ICE UNDER A FILM IN THE CASE OF A LAMINAR FLOW

Let us consider a laminar motion of a film. Taking into account relations (10), (14), and (15), we can represent the velocity of crystallization-front motion in the form

$$\frac{dz'}{dt} = \frac{k_0}{ut - z'}, \quad (23)$$

where

$$k_0 = \frac{\lambda \Delta T}{\rho[L_c - c(T_0 - T_\infty)]}. \quad (24)$$

In terms of the variable $h = ut - z'$, Eq. (23) assumes the form

$$\frac{dh}{dt} + \frac{k_0}{h} = u. \quad (25)$$

From the condition $dh/dt = 0$, we find the equilibrium film thickness. The result is

$$h_b = \frac{k_0}{u} = \frac{\lambda \Delta T}{qEV[L_c - c(T_0 - T_\infty)]}. \quad (26)$$

From (25) and (26), we obtain

$$\frac{dh}{dt} = u\left[1 - \frac{h_b}{h}\right]. \quad (27)$$

From (27), one can see that, if $h < h_b$, then $dh/dt < 0$, which means that the film thickness decreases; but if $h > h_b$, the film thickness increases. Thus, the h_b state is unstable, since an arbitrary small deviation of h from h_b

triggers changes in h that are directed away from the equilibrium value.

Solving Eq. (25), we arrive at

$$h - h_0 + h_b \ln\left|\frac{h-h_0}{h_0-h_b}\right| = ut. \quad (28)$$

Setting $h = 0$ in (28), we can find the time t_0 within which a film whose initial thickness satisfies the condition $h_0 < h_b$ disappears. The result is

$$t_0 = \frac{h_b \ln\left[\frac{h_b}{h_b - h_0}\right] - u}{u}. \quad (29)$$

Thus, the film disappears within the time t_0 if $h_0 < h_b$; in this case, drops crystallize without merging together, whereupon there arises a matt inhomogeneous structure of ice. But if $h_0 > h_b$ and if a viscous mode of film motion is realized (molecular mechanism of heat transfer), an increase in the film thickness leads to a transition to a turbulent mode (turbulent mechanism of heat transfer).

4. CRITICAL FILM THICKNESS

Let us now find the critical film thickness at which there occurs a transition from a viscous to a turbulent mode of film motion. From (9) and (26), we obtain

$$h_e h_b = k_0 \tau = \frac{\lambda}{\lambda_t} h^2. \quad (30)$$

From here, we find the critical film thickness h_c, which corresponds to the condition $\lambda_t = \lambda$,

$$h_c = \sqrt{k_0 \tau} = \sqrt{h_e h_b} = \sqrt{\frac{\lambda}{c\rho B}} \frac{\mathrm{Re}_a^{-0.1}}{V}, \quad (31)$$

that is, the critical film thickness h_c is the film thickness at which the coefficient of turbulent thermal conductivity is equal to the coefficient of molecular thermal conductivity. Although both heat-transfer mechanisms coexist in a film [7], we will assume, as is done in many problems of fluid dynamics [10], that the dominant mechanism is molecular for $h < h_c$ and turbulent for $h > h_c$.

As can be seen from (31), the critical film thickness is not constant—it is in inverse proportion to the airflow velocity. Substituting expression (31) for h_c into Eqs. (18) and (19), we obtain

$$\left[\frac{dh}{dt}\right]_c = u - \sqrt{k_0/\tau} = u[1 - \sqrt{h_b/h_e}]. \quad (32)$$

From (31) and (32), it follows that three different situations are possible: (i) $h_b < h_c < h_e$, in which case $(dh/dt)_c > 0$; (ii) $h_b = h_c = h_e$, in which case $(dh/dt)_c = 0$; and (iii) $h_b > h_c > h_e$, in which case $(dh/dt)_e < 0$. In the first case ($h_b < h < h_e$), the mode of film motion is laminar, with the result that the film thickness grows to the value h_e, at which there occurs a transition to turbulent mode. After that, the film thickness continues growing up to the value h_e. In the second case, the film thickness is constant, the mode of film motion being turbulent. In the third case, the film thickness decreases, irrespective of the mode of film motion.

Thus, there is a critical equilibrium-film-thickness value $h_{b.c}$ that corresponds to the equilibrium state $h_b = h_c = h_e$ and which can be determined only from experiments. Experiments revealed [1, 2] that $h_{b.c} = 1.3 \times 10^{-3}$ m for water.

Thus, we see that, if $h < h_{b.c}$, the mechanism of thermal conductivity is molecular, the film disappears, and drops crystallize without merging; as a result, there arises a matt inhomogeneous structure of ice. But if $h > h_{b.c}$, the mechanism of thermal conductivity is turbulent, in which case the film thickness decreases, tending to $h_{b.c}$; here, crystallization occurs under the film of constant thickness, this resulting in the formation of a transparent homogeneous structure of ice.

As can be seen from (18), (25), and (32), some critical velocity u_c of the film surface,

$$u_c = \left[\frac{qEV}{\rho}\right]_c = \frac{k_0}{h_{b.c}} = \frac{\lambda \Delta T}{\rho h_{b.c}[L_c - c(T_0 - T_\infty)]} \quad (33)$$

corresponds to the stable state $(dh/dt)_c = 0$.

If, in (33), $c(T_0 - T_\infty)$ is disregarded in respect to L_c, the transition from one structure of ice to the other is determined by the generalized Maclin parameter [13]

$$M = \frac{qEV}{\Delta T} = \frac{\lambda}{L h_{b.c}}. \quad (34)$$

Thus, both the Maclin parameter and the equilibrium critical film thickness characterize the transition from one structure of ice to the other. As to formula (34), it establishes a relationship between the parameters M and $h_{b.c}$.

For the water content in the flux, formula (33) makes it possible to find the critical value q_K determining the transition from one structure of ice to the other. We have

$$q_K = \frac{\lambda \Delta T}{h_{b.c} EV[L_c - c(T_0 - T_\infty)]}. \quad (35)$$

We refer to this quantity as Kachurin's critical water content. By way of example, we indicate that, at $E = 1$, $\Delta T = 1°\mathrm{C}$, and $V = 10^2$ m/s, Kachurin's critical water content is $q_K \cong 3 \times 10^{-4}$ kg/m^3.

With allowance for (11) and (14), the heat-balance equation at the crystallization front assumes the form

$$\rho[L_c - c(T_0 - T_\infty)]\frac{dz'}{dt} = \alpha(T_s - T_\infty) + L\beta(\rho_s - \rho_\infty). \quad (36)$$

Let us consider the particular case where there is no film at the plate surface and where drops crystallize without merging together. We then have $dz'/dt = dz''/dt$. Substituting (15) into (30), we arrive at

$$qEV[L_c - c(T_0 - T_\infty)] = \alpha(T_s - T_\infty) + L\beta(\rho_s - \rho_\infty). \quad (37)$$

From this relation, we find the critical water content q_c',

$$q_c' = \frac{\alpha(T_s - T_\infty) + L\beta(\rho_s - \rho_\infty)}{EV[L_c - c(T_0 - T_\infty)]}, \quad (38)$$

which determines the condition under which all drops that are deposited onto the plate surface crystallize. If $q > q_c'$, a film is formed on the plate surface. The condition $q > q_c'$ is not sufficient for the formation of a transparent homogeneous structure of ice, since, for $h < h_b$, the film at the plate surface is unstable and disappears with time. For the formation of a transparent homogeneous structure of ice, it is necessary that the film thickness be larger than the respective equilibrium value.

5. SEMIEMPIRICAL THEORY OF THE FORMATION OF A LAYERED STRUCTURE OF ICE

From the phenomenological theory developed in [1, 2] for describing the formation of a layered structure of ice, it follows that the layered structure of ice is determined by the parameter h_b, which is the equilibrium film thickness. At various values of h_b, results obtained by measuring the density ρ_i of ice layers on obstacles streamlined by a water-aerosol flow are well approximated by the formula [14, 15]

$$\rho_i = \rho_0\left(1 - \exp\left(-\frac{0.32}{h_b}\right)\right), \quad (39)$$

where ρ_0 is the density (in kg/m^3 units) of transparent homogeneous ice.

From (39), it follows that the formation of a transparent homogeneous structure of ice at small h_b and the formation of a matt inhomogeneous structure of ice at large h_b occur. The same conclusion can be drawn from formula (27). Indeed, we can see that, at small h_b, there is a high probability that $h_0 > h_b$, in which case the film thickness grows, according to (27); crystallization then proceeds under the film. As a result, there arises a transparent homogeneous structure of ice (wet growth). For large h_b ($h_b \longrightarrow \infty$), there is a high probability that $h_0 < h_b$. According to (27), the film thickness then decreases; the film disappears, and crystallization proceeds in the absence of a film. This leads to the formation of an inhomogeneous structure of ice. On the basis of this statistical treatment, the parenthetical expression on the right-hand side of (39),

$$P = 1 - \exp(-0.32/h_b), \quad (40)$$

can be interpreted as the empirical probability distribution for the formation of a transparent homogeneous structure of ice. Indeed, we see that, at $h_b = 0$, $P = 1$—that is, the formation of a transparent homogeneous structure of ice is expected here with the probability of $P = 1$. For $h_b \longrightarrow \infty$, we have $P = 0$; that is, the probability for the formation of a transparent homogeneous structure of ice is zero, which is equivalent to the probability $Q = 1 - P$ of unity for the formation of a matt inhomogeneous structure of ice. From the above interpretation, it follows that one can introduce the critical equilibrium-film-thickness value $h_{b.c}$ at which the formation of a transparent homogeneous structure of ice is expected with a 90% probability, $P = 0.9$. From (40), we obtain

$$h_{b.c} = \frac{0.32}{\ln 10} \approx 1.3 \times 10^{-3}\ (\text{m}). \quad (41)$$

In other words, the meaning of $h_{b.c}$ is as follows: according to (27), one can expect $h_0 > h_{b.c}$ with a probability of 90%. This will lead to the formation of a transparent homogeneous structure of ice. In the statistical interpretation, the layered structure of ice does not depend on the motion of the film. The equilibrium film thickness h_b can be considered as a criterion that specifies the transition from one structure of ice to the other. It should be noted in addition that, according to [7, 10], there is no clear-cut boundary between the laminar and the turbulent motion of the film. In the region of the critical Reynolds number, there is intermittency, in which case a laminar mode gives way to a turbulent mode, and vice versa. The intermittency phenomenon is quantitatively characterized by the parameter γ that determines the fraction of time within which the motion in question is of a turbulent character. From this point of view, the critical film thickness determining the transition from laminar to turbulent motion should be interpreted as a statistical property. For $h > h_b$, the motion can be either laminar or turbulent. In the case of a laminar motion, the film thickness grows ($h > h_c$) up to some value at which drops are stripped off the film surface. But in the case of a turbulent motion, the film thickness tends to the constant value h_e, in which case an arbitrary great water content can be absorbed.

Thus, only within the statistical approach is it possible to explain satisfactorily the role of the critical equilibrium film thickness $h_{b.c}$ [see also Eq. (39)]. The phenomenological approach alone discloses the role of the parameter h_b—it is impossible to assess $h_{b.c}$ within this approach.

6. CONCLUSIONS

For a film to be preserved at the plate surface steadily, the film thickness must be larger than the critical equilibrium thickness $h_{b.c}$, in which case a transparent homogeneous structure of ice is formed. Otherwise, the film at the plate surface proves to be unstable and disappears with time. This leads to the formation of a matt inhomogeneous structure of ice.

In just the same way as was done in the present article above for a plate and in [6] for the formation of a layered structure of ice at the surface of a sphere (growth of hailstones), one can construct a theory that would describe the formation of a layered structure of ice at the surface of a cylinder (icing of wires). For this, it is necessary to take into account the corresponding expression for the drag resistance c_f [10], the corresponding shape coefficient ($k_f = 2$ for a cylinder), and the corresponding Nusselt and Sherwood numbers [7, 10] (Nu and Sh, respectively). As to the mechanism responsible for the formation of a layered structure of ice, it is identical for a plate, a sphere, and a cylinder.

Thus, Kachurin's theory [1] has been extended in this study with allowance for heat and mass transfer at the film surface. Also, the ways for consistently applying it to surfaces of various shapes have been indicated.

REFERENCES

1. L. G. Kachurin, Izv. Akad. Nauk SSSR, Ser. Geofiz., No. 6, 823 (1962).
2. L. G. Kachurin and V. G. Morachevskiĭ, *Kinetics of the Phase Transitions of Water in Atmosphere* (Leningr. Gos. Univ., Leningrad, 1965), p. 144 [in Russian].
3. R. G. Zakinyan, Tr. SF VGI (Stavropol Filial of Alpine Geophysical Institute), No. 1, 122 (1993).
4. R. G. Zakinyan, Tr. SF VGI (Stavropol Filial of Alpine Geophysical Institute), No. 1, 130 (1993).
5. R. G. Zakinyan, Tr. SF VGI (Stavropol Filial of Alpine Geophysical Institute), No. 1, 140 (1993).
6. R. G. Zakinyan, Inzh.-Fiz. Zh. **76** (2), 30 (2003).
7. L. G. Loĭtsyanskiĭ, *Mechanics of Liquids and Gases* (Nauka, Moscow, 1978) [in Russian].
8. J. O. Hinze, *Turbulence. An Introduction to Its Mechanism and Theory* (McGraw-Hill, New York, 1959; Fizmatlit, Moscow, 1963).
9. V. G. Levich, *Physicochemical Hydrodynamics* (Fizmatgiz, Moscow, 1959) [in Russian].
10. H. Schlichting, *Boundary Layer Theory*, 6th ed. (McGraw-Hill, New York, 1968; Nauka, Moscow, 1974).
11. L. Prandtl, *Strömungslehre* (Vieweg, Braunschweig, 1960; PKhD, Moscow, 2002) [translated from German].
12. A. N. Tikhonov and A. A. Samarskiĭ, *Equations of Mathematical Physics*, 4th ed. (Nauka, Moscow, 1977; Pergamon Press, Oxford, 1964).
13. W. C. Maclin, Q. J. R. Meteorol. Soc. **94** (401), 73 (1968).
14. L. G. Kachurin and L. I. Gashin, Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana **4**, 93 (1968).
15. L. G. Kachurin, *Physical Principles of Action on Atmospheric Processes* (Gidrometeoizdat, Leningrad, 1990) [in Russian].

Translated by A. Isaakyan

Technical Physics, Vol. 49, No. 9, 2004, pp. 1116–1125. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 15–22.

GASES AND LIQUIDS

Initiation of Combustion of a Methane–Air Mixture in a Supersonic Flow Behind a Shock Wave during Laser Excitation of O_2 Molecules

A. M. Starik and N. S. Titova

Federal State Unitary Enterprise "Baranov Central Institute of Aviation Motors," Moscow, 111116 Russia
e-mail: star@ciam.ru

Received January 8, 2004

Abstract—The possibility of initiating detonation of CH_4 + air in a supersonic flow behind an oblique shock wave under the exposure of the mixture to laser radiation with wavelengths λ_I = 1.268 μm and 762 nm is analyzed. It is shown that this irradiation leads to excitation of O_2 molecules to the $a^1\Delta_g$ and $b^1\Sigma_g^+$ states, which intensifies the chain mechanism of combustion of CH_4/O_2 (air) mixtures. Even for a small value of the laser radiation energy absorbed by an O_2 molecule (~0.05–0.1 eV), detonation mode of combustion in a poorly inflammable mixture such as CH_4/air can be realized at a distance of only 1 m from the primary shock wave front for relatively small values of temperature (~1100 K) behind the front under atmospheric pressure. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Considerable attention in recent years toward analysis of detonation burning of various mixtures [1–6] is not only due to the possibility of studying the kinetics of combustion in "purely" gasdynamic conditions, but also due to prospects of designing hypersonic oblique detonation wave engines and pulse detonation engines [7–9]. The key problem in realization of detonation mode of combustion in a hypersonic flow is the reduction of the length of the induction zone, which is too large (>25 m) even for hydrogen–air mixtures at temperatures T_1 < 850 K behind the shock wave front.

It was shown earlier that preliminary excitation of H_2 and N_2 molecular vibrations [10] or O_2 molecules to the $a^1\Delta_g$ electronic state [11] makes it possible to initiate a detonation wave in a supersonic flow of an H_2/O_2 mixture at a distance of ~1 m from the primary shock wave front even at moderate temperatures T_1 = 600 K behind the front. This is due to intensification of the chain combustion mechanism. Since molecular oxygen is an active oxidizer not only for H_2, but also for various hydrocarbons (including methane), it would be interesting to determine how the excitation of O_2 molecules to the $a^1\Delta_g$ or $b^1\Sigma_g^+$ state affects the length of the induction zone during the combustion in the supersonic flow of CH_4/O_2 (air) mixtures, induced by an oblique shock wave.

One of the method for obtaining excited $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules is the exposure of the reactive mixture to laser radiation with a wavelength of λ_I = 1.268 μm and 762 nm, respectively. In our recent publication [12], it was shown that the action of laser radiation with λ_I = 762 nm on a supersonic flow of the stoichiometric mixture H_2 + O_2 in front of an oblique shock wave makes it possible to realize the combustion at distances shorter than 1 m from the wavefront even for small values of laser radiation energy flux (~1 J/cm^2). It should be noted that earlier [13], it was proposed that detonation should be initiated in a supersonic flow due to thermal action of laser radiation. However, the efficiency of this method is low (see below).

This study is devoted to analysis of the kinetic mechanisms leading to initiation of combustion in a supersonic flow of a CH_4/O_2 (air) mixture by exciting O_2 molecules by radiation with λ_I = 1.268 μm and 762 nm.

MAIN ASSUMPTIONS AND FORMULATION OF THE PROBLEM

Let us consider schematically a flow with a stationary shock wave (Fig. 1). Here, a homogeneous mixture of CH_4 + air flowing at supersonic velocity is subjected to the action of radiation of constant intensity I_0 over an interval δ in front of the shock wave. The radiation frequency ν_I is in resonance with the frequency of the bound–bound electron transition $m(e', v', j', K') \longrightarrow n(e'', v'', j'', K'')$ in the O_2 molecule, where $e' = X^3\Sigma_g^-$, $e'' = a^1\Delta_g$ or $b^1\Sigma_g^+$, v' and v'' are the vibrational and j', K' and j'', K'' the rotational quantum numbers in states e' and e'', respectively. The angle of inclination of the

1063-7842/04/4909-1116$26.00 © 2004 MAIK "Nauka/Interperiodica"

wavefront to the velocity vector u_0 of the unperturbed flow is $\beta \le 30°$. In this case, the velocity of the gas behind the front remains supersonic.

We will treat the electron-excited molecules $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ as individual chemical species with the corresponding enthalpy of formation and assume that the vibrational, rotational, and translational degrees of freedom of molecules in the mixture are in thermodynamic equilibrium, which is not violated in the case of radiation-induced transitions and chemical reactions. We assume that the gas is inviscid and nonheat conducting. Suppose that $\delta \ll L_\nu$, where L_ν is the absorption length for laser radiation. Under these assumptions, we can write the system of equations describing the physicochemical processes in the radiation zone and behind the shock wave front in the form

$$u\frac{dN_i}{dx} = Q_{Ii} + Q_{ci} + Q_{si}, \tag{1}$$

$$u\frac{du}{dx} + \frac{1}{\rho}\frac{dP}{dx} = 0, \tag{2}$$

$$\frac{dH}{dx} + u\frac{du}{dx} = \frac{k_\nu I_0}{\rho u},$$

$$H = \sum_{i=1}^{M}\frac{h_{0i}}{\mu}\gamma_i + C_pT,$$

$$C_p = \frac{R}{\mu}\left(\frac{5}{2} + \sum_{i=1}^{S}C_R^i\gamma_i + \sum_{i=1}^{S}C_v^i\gamma_i\right),$$

$$\mu = \sum_{i=1}^{M}\mu_i\gamma_i, \quad P = \frac{\rho RT}{\mu},$$

$$C_v^i = \sum_{j=1}^{L}\left(\frac{\theta_{ij}}{T}\right)^2\frac{\exp(\theta_{ij}/T)}{[\exp(\theta_{ij}/T)-1]^2}, \quad \gamma_i = \frac{N_i}{N},$$

$$N = \sum_{i=1}^{M}N_i,$$

$$Q_{ci} = \sum_{q=1}^{M_1}S_{iq}, \quad S_{iq} = (\alpha_{iq}^- - \alpha_{iq}^+)[R_q^+ - R_q^-],$$

$$R_q^{+(-)} = k_{+(-)q}\prod_{j=1}^{n_q^{+(-)}}N_j^{\alpha_{iq}^{+(-)}},$$

$$Q_{Ii} = l_{iI}W_I\left(\frac{g_n}{g_m}N_m - N_n\right), \quad W_I = \sigma_{mn}I/h\nu_I,$$

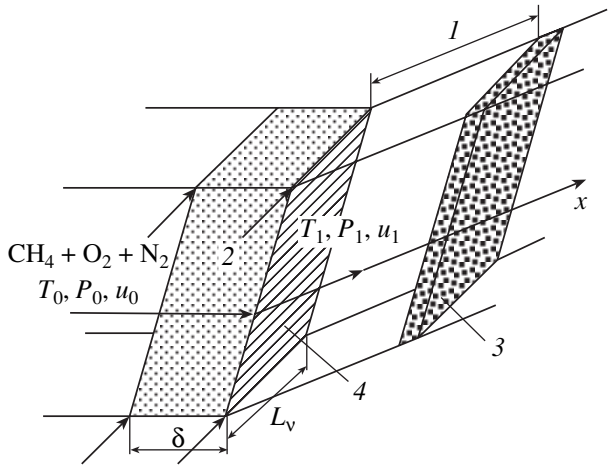


Fig. 1. Schematic diagram of a supersonic flow in which combustion is initiated by laser radiation: (*1*) induction zone, (*2*) laser radiation, (*3*) combustion front, (*4*) shock wave front.

$$\sigma_{mn} = \frac{\lambda_{mn}^2}{4\pi b_D}A_{mn}\sqrt{\frac{\ln 2}{\pi}}H(x,a), \quad Q_{si} = \sum_q r_{iq}^s\sum_j A_{qj}^s N_j,$$

$$k_\nu = \sigma_{mn}\left(\frac{g_n}{g_m}N_m - N_n\right), \quad N_m = N_1\varphi_m,$$

$$N_n = N_l\varphi_n, \quad l = 2 \text{ or } 3,$$

$$\varphi_m = \frac{g_mB_{v'}}{kT}\frac{\exp(-\theta_1 v'/T)}{1-\exp(-\theta_1/T)}\exp\left(-\frac{E_{j'}}{kT}\right),$$

$$\varphi_n = \frac{g_nB_{v''}}{kT}\frac{\exp(-\theta_l v''/T)}{1-\exp(-\theta_l/T)}\exp\left(-\frac{E_{j''}}{kT}\right). \tag{3}$$

Here, P, ρ, T, and u are the gas pressure, density, temperature, and velocity; N_i is the density of molecules (atoms) of the ith species (i = 1, 2, 3 correspond to $O_2(X^3\Sigma_g^-)$, $O_2(a^1\Delta_g)$, $O_2(b^1\Sigma_g^+)$; μ_i is their molecular mass; h_{0i} is the enthalpy of formation of the ith component at T = 298 K; M is the number of atomic and molecular components in the mixture; S is the number of molecular components alone; $C_R^i = 1$ for components from linear molecules and $C_R^i = 1.5$ for components of nonlinear molecules; θ_{ij} is the characteristic vibrational temperature of the jth mode for the ith component ($j = 1, \ldots, L$); M_1 is the number of reactions leading to the formation (disappearance) of the ith component; α_{iq}^+ and α_{iq}^- are the stoichiometric coefficients for the qth reaction; $n_q^{+(-)}$ is the number of components participating in the direct (+) and inverse (–) reaction; $k_{+(-)q}$ are the rate constants of these reactions; R is the gas constant; h is the Plank constant; k is the Boltzmann

constant, l_{il} is the number of quanta lost (acquired) by the ith component for induced transitions; r_{iq}^s is the same for spontaneous transitions; N_m and N_n are the numbers of molecules in the lower and upper states of the absorbing transition $m \longrightarrow n$; g_m and g_n are the degeneracies of these transitions, λ_{mn} is the wavelength corresponding to the center of the spectral line of the absorbing transition; $A_{mn}(A_{qj}^s)$ is the Einstein coefficient; b_D is the Doppler width of the spectral line of the $m \longrightarrow n$ transition; $H(x, a)$ is the Voigt function; B_v is the rotational constant of the O_2 molecule in state v ($v' \in m$, $v'' \in n$); and $E_{j'}$ and $E_{j''}$ are the rotational energies of the O_2 molecule in the m and n states. Their values were calculated taking into account the splitting of the j'th level in state $X^3\Sigma_g^-$ into three components with $j' = K' + 1$, $j' = K'$, and $j' = K' - 1$.

In numerical integration of system (1)–(3), the entire calculated domain splits into two subdomains. The first subdomain corresponds to the zone of exposure to resonance radiation in front of the oblique shock wave (its length $x_0 \le \delta < x_1$), while the second corresponds to the reaction zone ($x > x_1$). The boundary conditions for the system of equations (1)–(3) for $x = x_1$ are the parameters behind the shock wave front (these parameters will henceforth be marked by index 1), which can be determined by solving the following system of equations [14]:

$$\lambda_{n1} = \lambda_{n0}^{-1}, \quad H_1 - H_0 = \frac{1}{2}\frac{P_0}{\rho_0}\left(\frac{P_1}{P_0} - 1\right)\left(\frac{\rho_0}{\rho_1} + 1\right),$$

$$-\chi_{0e}M_{n0}^2 = \left(\frac{P_1}{P_0} - 1\right)\left(\frac{\rho_0}{\rho_1} - 1\right),$$

$$\lambda_n = u_n/\sqrt{\frac{2\chi_e RT}{(\chi_e + 1)\mu}}, \quad M_n = u_n/\sqrt{\chi_e \frac{R}{\mu}T},$$

$$\chi_e = 1 + \left\{\frac{\mu}{R}C_p - 1\right\}^{-1}, \quad u_1 = \sqrt{u_{n1}^2 + u_{\tau 1}^2}.$$

Here, u_n is the normal and u_τ is the tangential component of the flow velocity relative to the front and subscript 0 corresponds to parameters in front of the shock wave. Numerical integration of Eqs. (1)–(3) was carried out in the same way as in [10–12] using the implicit difference scheme of the second order approximation.

KINETIC MODEL

It is well known that the description of ignition and combustion of CH_4/O_2 (air) mixtures in a wide range of initial temperatures and pressures requires complex kinetic schemes [15–17]. We used the scheme proposed in [17] as the basis for describing the reaction of methane with air in the bulk; this scheme contained 433 reversible reactions involving 58 components. However, not all of these reactions play a significant role in initiation of combustion behind the shock wave. To determine the minimal set of processes ensuring a correct description of the dynamics of ignition and heat release in the CH_4/air mixture behind an oblique shock wave in the absence of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules, an analysis of possible reduction of the total kinetic scheme [17] was performed.

It was proved earlier that, for the ignition of stoichiometric and fuel-depleted CH_4/O_2 mixtures in a closed adiabatic reactor, it is possible to obtain correct values of the induction period, τ_{in}, and the final gas temperature, T_c, by using a reduced (relative to the complete scheme [17]) model including 211 reversible reactions involving 35 components (CO_x (x = 1, 2), HO_x, H_2O_x, H_x, O_y (y = 1, 2, 3), C_x, CH_q (q = 1, …, 4), C_2H_z (z = 1, …, 6), CH_qO_x, and $C_2H_qO_x$). In the presence of N_2 molecules in the initial mixture (CH_4/air), these reactions should be supplemented with those involving N, N_2, NO, and NO_2 (reactions 271–293 from [17]). The inclusion of processes with other N-containing species does not lead to a noticeable change in τ_{in} and T_c. For this reason, the given scheme (scheme 1) was treated as basic. In addition, we considered a scheme from which reactions with participation of C_2, CH_2CO, CH_3CO, C_2HO, CH_3CHO, and $CH_3O_2CH_3$ were excluded (as compared to scheme 1), as well as scheme 3, consisting of 154 reactions with 29 components (excluding additional reactions with the participation of CH_2OH, CH_3OH, CH_3O_2, and CH_3O_2H). The results of calculating the induction zone length L_{in} and combustion zone L_c (the quantity L_{in} was defined as the distance from the front on which the maximal value of gradient dT/dx is attained and L_c was defined as the distance at which $T = 0.99T_c$), the values of T_c, P_c, and M_c at the end of the reaction zone for the stoichiometric CH_4/air mixture ($CH_4/O_2/N_2$ = 0.5/1/3.76) with $P_0 = 10^4$ Pa and T_0 = 300 K for various values of M_0 and the angle of front inclination $\beta = 30°$ proved that reduced scheme 2 provides a satisfactory accuracy in L_{in}, L_c, T_c, M_c, and P_c in the range of variation of the initial flow parameters under investigation ($P_0 = 10^2$–10^5 Pa, T_0 = 300 K, and M_0 = 6–10).

This scheme was supplemented with reactions involving $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules. All these reactions are compiled in the table. Since the reaction with the participation of electron-excited O_2 molecules occur with an energy barrier lower than in reactions with $O_2(X^3\Sigma_g^-)$ [12, 18], we supplemented the scheme with the reactions of N_2 with $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ as well as those with the participation of these molecules and N_2O, in which chemically active O atoms and O_3

Reactions with participation of excited molecules $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ included in the kinetic scheme for calculating the combustion of methane in air

No.	Reaction	No.	Reaction
1	$O_2(a^1\Delta_g) + M = 2O + M$	28	$O_3 + HO_2 = OH + O_2(X^3\Sigma_g^-) + O_2(b^1\Sigma_g^+)$
2	$O_2(b^1\Sigma_g^+) + M = 2O + M$	29	$O_3 + O_2(a^1\Delta_g) = 2O_2(X^3\Sigma_g^-) + O$
3	$O_2(a^1\Delta_g) + H = OH + O$	30	$O_3 + O_2(b^1\Sigma_g^+) = 2O_2(X^3\Sigma_g^-) + O$
4	$O_2(b^1\Sigma_g^+) + H = OH + O$	31	$2O_2(a^1\Delta_g) = O_2(X^3\Sigma_g^-) + O_2$
5	$H_2 + O_2(a^1\Delta_g) = 2OH$	32	$CH_2O + O_2(a^1\Delta_g) = HO_2 + HCO$
6	$H_2 + O_2(b^1\Sigma_g^+) = 2OH$	33	$CH_2O + O_2(b^1\Sigma_g^+) = HO_2 + HCO$
7	$HO_2 + M = O_2(a^1\Delta_g) + H + M$	34	$HO_2 + CO = HCO + O_2(a^1\Delta_g)$
8	$HO_2 + M = O_2(b^1\Sigma_g^+) + H + M$	35	$HO_2 + CO = HCO + O_2(b^1\Sigma_g^+)$
9	$H_2 + O_2(a^1\Delta_g) = H + HO_2$	36	$CH_4 + O_2(a^1\Delta_g) = CH_3 + HO_2$
10	$H_2 + O_2(b^1\Sigma_g^+) = H + HO_2$	37	$CH_4 + O_2(b^1\Sigma_g^+) = CH_3 + HO_2$
11	$H_2O + O_2(a^1\Delta_g) = OH + HO_2$	38	$CH_3 + O_2(a^1\Delta_g) = CH_3O + O$
12	$H_2O + O_2(b^1\Sigma_g^+) = OH + HO_2$	39	$CH_3 + O_2(b^1\Sigma_g^+) = CH_3O + O$
13	$OH + O_2(a^1\Delta_g) = O + HO_2$	40	$CH_3O + O_2(a^1\Delta_g) = CH_2O + HO_2$
14	$OH + O_2(b^1\Sigma_g^+) = O + HO_2$	41	$CH_3O + O_2(b^1\Sigma_g^+) = CH_2O + HO_2$
15	$2HO_2 = H_2O_2 + O_2(a^1\Delta_g)$	42	$CH_3 + O_2(a^1\Delta_g) = CH_2O + OH$
16	$2HO_2 = H_2O_2 + O_2(b^1\Sigma_g^+)$	43	$CH_3 + O_2(b^1\Sigma_g^+) = CH_2O + OH$
17	$H_2O + O_2(a^1\Delta_g) = H_2O_2 + O$	44	$CO + O_2(a^1\Delta_g) = CO_2 + O$
18	$H_2O + O_2(b^1\Sigma_g^+) = H_2O_2 + O$	45	$CO + O_2(b^1\Sigma_g^+) = CO_2 + O$
19	$O_3 + M = O_2(a^1\Delta_g) + O + M$	46	$CH + O_2(a^1\Delta_g) = CO + OH$
20	$O_3 + M = O_2(b^1\Sigma_g^+) + O + M$	47	$CH + O_2(b^1\Sigma_g^+) = CO + OH$
21	$O_3 + H = OH + O_2(a^1\Delta_g)$	48	$HCO + O = CH + O_2(a^1\Delta_g)$
22	$O_3 + H = OH + O_2(b^1\Sigma_g^+)$	49	$HCO + O = CH + O_2(b^1\Sigma_g^+)$
23	$O_3 + O = O_2(X^3\Sigma_g^-) + O_2(a^1\Delta_g)$	50	$CH_2OH + O_2(a^1\Delta_g) = CH_2O + HO_2$
24	$O_3 + O = O_2(X^3\Sigma_g^-) + O_2(b^1\Sigma_g^+)$	51	$CH_2OH + O_2(b^1\Sigma_g^+) = CH_2O + HO_2$
25	$O_3 + OH = HO_2 + O_2(a^1\Delta_g)$	52	$CH_3O_2 + OH = CH_3OH + O_2(a^1\Delta_g)$
26	$O_3 + OH = HO_2 + O_2(b^1\Sigma_g^+)$	53	$CH_3O_2 + OH = CH_3OH + O_2(b^1\Sigma_g^+)$
27	$O_3 + HO_2 = OH + O_2(X^3\Sigma_g^-) + O_2(a^1\Delta_g)$	54	$CH_3O_2 + CH_3O_2 = CH_3OH + CH_2O + O_2(a^1\Delta_g)$

Table. (Contd.)

No.	Reaction	No.	Reaction
55	$CH_3O_2 + CH_3O_2 = CH_3OH + CH_2O + O_2(b^1\Sigma_g^+)$	74	$C_2H_2 + O_2(a^1\Delta_g) = C_2H + HO_2$
56	$CH_3 + O_2(a^1\Delta_g) = CH_3O_2$	75	$C_2H_2 + O_2(b^1\Sigma_g^+) = C_2H + HO_2$
57	$CH_3 + O_2(b^1\Sigma_g^+) = CH_3O_2$	76	$C_2H + O_2(a^1\Delta_g) = CO + HCO$
58	$CH_3O_2 + O = CH_3O + O_2(a^1\Delta_g)$	77	$C_2H + O_2(b^1\Sigma_g^+) = CO + HCO$
59	$CH_3O_2 + O = CH_3O + O_2(b^1\Sigma_g^+)$	78	$N + O_2(a^1\Delta_g) = O + NO$
60	$CH_3O_2 + HO_2 = CH_3O_2H + O_2(a^1\Delta_g)$	79	$N + O_2(b^1\Sigma_g^+) = O + NO$
61	$CH_3O_2 + HO_2 = CH_3O_2H + O_2(b^1\Sigma_g^+)$	80	$NO + O_2(a^1\Delta_g) = O + NO_2$
62	$CH_3O_2 + CH_3O_2 = CH_3O + CH_3O + O_2(a^1\Delta_g)$	81	$NO + O_2(b^1\Sigma_g^+) = O + NO_2$
63	$CH_3O_2 + CH_3O_2 = CH_3O + CH_3O + O_2(b^1\Sigma_g^+)$	82	$NO + NO + O_2(a^1\Delta_g) = NO_2 + NO_2$
64	$C_2H_6 + O_2(a^1\Delta_g) = C_2H_5 + HO_2$	83	$NO + NO + O_2(b^1\Sigma_g^+) = NO_2 + NO_2$
65	$C_2H_6 + O_2(b^1\Sigma_g^+) = C_2H_5 + HO_2$	84	$O_3 + NO = NO_2 + O_2(a^1\Delta_g)$
66	$C_2H_5 + O_2(a^1\Delta_g) = C_2H_4 + HO_2$	85	$O_3 + NO = NO_2 + O_2(b^1\Sigma_g^+)$
67	$C_2H_5 + O_2(b^1\Sigma_g^+) = C_2H_4 + HO_2$	86	$N_2 + O_2(a^1\Delta_g) = N_2O + O$
68	$C_2H_4 + O_2(a^1\Delta_g) = C_2H_3 + HO_2$	87	$N_2 + O_2(b^1\Sigma_g^+) = N_2O + O$
69	$C_2H_4 + O_2(b^1\Sigma_g^+) = C_2H_3 + HO_2$	88	$O_3 + N_2 = N_2O + O_2(a^1\Delta_g)$
70	$C_2H_3 + O_2(a^1\Delta_g) = C_2H_2 + HO_2$	89	$O_3 + N_2 = N_2O + O_2(b^1\Sigma_g^+)$
71	$C_2H_3 + O_2(b^1\Sigma_g^+) = C_2H_2 + HO_2$	90	$O_2(a^1\Delta_g) + M = O_2(X^3\Sigma_g^-) + M$
72	$C_2H_2 + O_2(a^1\Delta_g) = HCO + HCO$	91	$O_2(b^1\Sigma_g^+) + M = O_2(a^1\Delta_g) + M$
73	$C_2H_2 + O_2(b^1\Sigma_g^+) = HCO + HCO$		

molecules are formed (in the presence of unexcited O_2 molecules, the rate of formation of O and O_3 in these reactions is low and does not affects the ignition process). In addition to chemical reactions with excited O_2 molecules, the model also included deexcitation processes for $a^1\Delta_g$ and $b^1\Sigma_g^+$ states. The rate constants of chemical reactions with the participation of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules were determined using the technique described in detail in [12], where the temperature dependences $k_{+(-)q}(T)$ are presented for reactions in the system H_2–O_2–$O_2(a^1\Delta_g)$–$O_2(b^1\Sigma_g^+)$. In particular, for the endothermic reactions nos. 32–47, 50, 51, 64–83, and 86, 87, the activation energy of the processes involving electronically excited O_2 molecules was determined by the formula

$$E_a^e = \frac{1}{2}\left(\sqrt{(\Delta H + E_e)^2 + 4E_a^0(\Delta H + E_a^0)} - (\Delta H + E_e)\right),$$

where ΔH is the thermal effect of the reaction, E_a^0 is the activation energy of the reaction with nonexcited molecules, and E_e is the energy of an electronically excited molecule.

The reaction rate constant itself was calculated from the known relation

$$k_q(T) = A_q T^{n_q} \exp(-E^e_{aq}/T).$$

Here, A_q is the coefficient in the Arrhenius dependence of the rate constant of a reaction with the participation of a nonexcited molecule and n_q is a power coefficient.

As in [12], for barrierless reactions in which an O_2 molecule appears in the states $X^3\Sigma_g^-$, $a^1\Delta_g$, and $b^1\Sigma_g^+$ (reactions nos. 52–55, 58–61, 84, 85, 88, and 89), we assumed that the probabilities of formation of $O_2(X^3\Sigma_g^-)$, $O_2(a^1\Delta_g)$, and $O_2(b^1\Sigma_g^+)$ are proportional to the degeneracy of these states: $q_X = 0.5$, $q_a = 0.33$, and $q_b = 0.17$. The rate constants to overall processes were borrowed from [17]. Calculating the rate constants of direct reactions 48, 49, 62, and 63, we increased the energy barrier by the energy of the corresponding electron state ($E^e_a = E^0_a + E_e$).

The rate constants of deexcitation (electron-translational *E–T* relaxation) of states $a^1\Delta_g$ and $b^1\Sigma_g^+$ of the O_2 molecule (reactions nos. 90 and 91) were taken from [12] for M = O, O_3, O_2, H_2, H_2O, HO_2, and H_2O_2 and from [18, 19] for M = CO_2, CO, and N_2. For M = H, C, and N, we assumed that deexcitation occurs with the same probability as for M = O; for M = CH_q, C_2H_q, CH_qO, CH_3O_2, and CH_3O_2H the same as for M = H_2O; for M = NO, NO_2, and N_2O, the same as for M = N_2. The rate constants of reactions occurring in the inverse direction were calculated on the basis of the principle of detailed balancing.

INITIATION OF COMBUSTION OF A CH_4/O_2 (AIR) MIXTURE IN A SUPERSONIC FLOW UNDER THE ACTION OF RADIATION

We will carry out specific analysis for a stoichiometric CH_4/air mixture ($CH_4/O_2/N_2 = 0.5/1/3.76$) for the case of radiation absorption in transitions with $v' = v'' = 0$, $j' = 9$, $j'' = K' = K'' = 8$ (for such rotational quantum numbers, at $T_0 = 300$ K, the absorption coefficient for both transitions under investigation, $X^3\Sigma_g^- \longrightarrow a^1\Delta_g$ and $X^3\Sigma_g^- \longrightarrow b^1\Sigma_g^+$, has the maximal value). The wavelengths corresponding to the centers of the lines for these transitions are 1.268 μm and 762 nm, while the Einstein coefficients are 2.58×10^{-4} s^{-1} and 8.5×10^{-2} s^{-1}, respectively. Calculating the Voigt functions $H(x, a)$, we assumed, following [12], that the coefficients of collision broadening of the spectral absorption line are proportional to the gas kinetic cross sections of the collision of an O_2 molecule with the *M*th partner (in the reaction region, $M = CH_4$, O_2, N_2).

At low temperatures of the gas in front of the shock wave ($T_0 = 300$ K), the chemical reaction rates in the irradiation zone are substantially lower than the rate of induced transitions and the *E–T* relaxation rate for states $a^1\Delta_g$ and $b^1\Sigma_g^+$. In this case, the concentration of molecules $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ in the irradiation zone is determined by the relation between the induced transition time $\tau_I = W_I^{-1}$, the *E–T* relaxation time, and the time of pulse duration $\tau_p = \delta/u_0$. For the flow parameters under consideration ($P_0 = 10^2$–10^4 Pa and $T_0 = 300$ K), the absorption zone length $L_\nu = k_\nu^{-1}$ amounts to 2×10^4–600 m for radiation with $\lambda_I = 1.268$ μm and 320–10 m for radiation with $\lambda_I = 762$ nm. For this reason, we always have $\delta \ll L_\nu$ and the irradiation of the gas in front of the shock wave can be performed by multiple transverse scanning by a laser beam with a radius of 0.2–1 cm to attain the required value of the radiant energy absorbed by the gas.

As the gas passes through the front of the shock wave, its temperature and pressure increase, leading to an increase in the rate of chemical reactions and to the ignition of the mixture at a certain distance from the front. The presence of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules in the mixture sharply accelerates this process. This is illustrated in Fig. 2, showing the variation of the mole fraction (γ_i) of the main components responsible for the development of the chain mechanism of ignition of the CH_4/air mixture in the irradiation zone as well as behind the shock wave front (Fig. 2a) in the absence of radiation and (Figs. 2b, 2c) under the action of radiation with $\lambda_I = 1.268$ μm and 762 nm, respectively, over the length $\delta = 50$ cm for $I_0 = 10$ kW/cm^2, $M_0 = 8$, and $\beta = 30°$. It can be seen that, even at a moderate temperature behind the front ($T_1 = 1102$ K), the action of radiation with $\lambda_I = 762$ nm leads to the ignition of the CH_4/air mixture at a distance of 8.8 m from the front even at a relatively low gas pressure of $P_1 = 1.9 \times 10^5$ Pa and at a moderate value of radiation energy density delivered to the gas, $E_{in} = I_0\tau_p = 1.76$ J/cm^2 (in the absence of irradiation, $L_{in} = 230$ m; i.e., the ignition of the CH_4/air mixture does not take place under the given conditions). Under the action of radiation with $\lambda_I = 1.268$ μm, the induction zone length L_{in} is substantially larger and amounts to 170 m. Such a difference can be explained by the fact that the radiation energy $E_s = k_\nu I_0 \tau_p / N_1$ absorbed by an O_2 molecule with the excited state $a^1\Delta_g$ is much lower ($E_s = 1 \times 10^{-4}$ eV/molecule) than the energy absorbed with the excitation of $b^1\Sigma_g^+$ ($E_s = 3.8 \times 10^{-2}$ eV/molecule).

It will be shown below that, for the same value of E_s, the value of L_{in} in the case of radiation with $\lambda_I = 1.268$ μm is even slightly smaller than for radiation

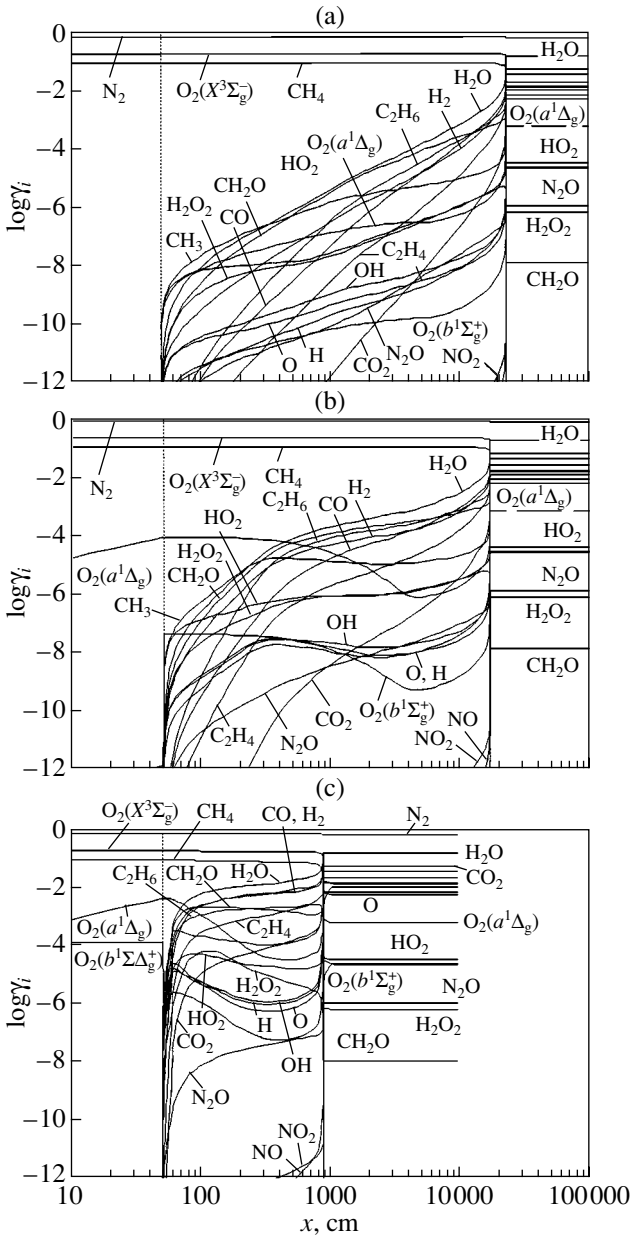


Fig. 2. Variation of concentration (mole fraction) of the components along the flow with $M_0 = 8$ and $\beta = 30°$ in the mixture $CH_4/O_2/N_2 = 0.5/1/3.76$ with $P_0 = 10^4$ Pa and $T_0 = 300$ K (a) in the absence of irradiation and (b, c) under the action of radiation with $\lambda_I = 1.268$ μm and 762 nm, respectively, for $I_0 = 10$ kW/cm^2 and $\delta = 50$ cm. The vertical dashed line corresponds to the position of the shock wave front.

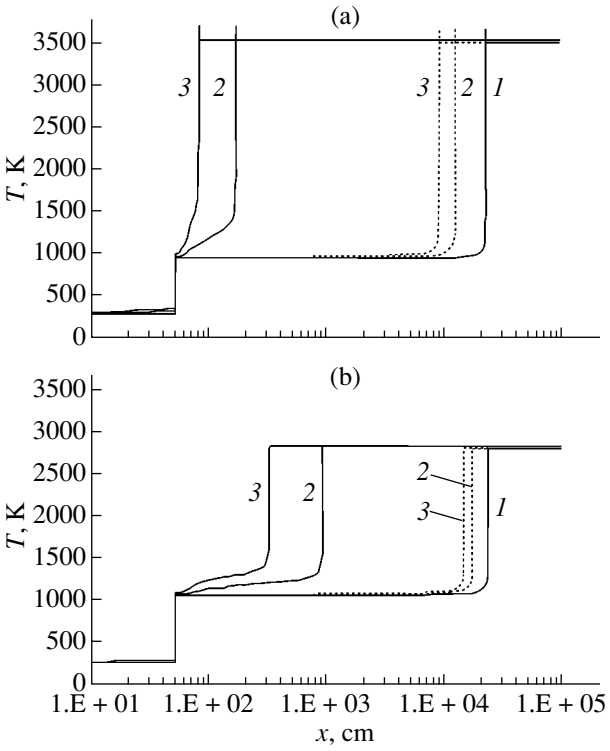


Fig. 3. Variation of the gas temperature in the irradiation zone behind the shock wave front ($M_0 = 8$, $\beta = 30°$) under the action radiation with $\lambda_I = 1.268$ μm (dotted curves) and 762 nm (solid curves) with $I_0 = 0$, 10, and 20 kW/cm^2 (*1*–*3*), $\delta = 50$ cm on the $CH_4/O_2 = 0.5/1$ (a) and $CH_4/O_2/N_2 = 0.5/1/3.76$ (b) mixtures at $T_0 = 300$ K, $P_0 = 10^4$ Pa.

with $\lambda_I = 762$ nm. This is due to the fact that the rate of relaxation of excited $O_2(b^1\Sigma_g^+)$ molecules via channel no. 91 (here and below, the numeration of reactions corresponds to that in the table) with the formation of $O_2(a^1\Delta_g)$ is much higher than the rate of relaxation of $O_2(a^1\Delta_g)$ molecules to the ground electron state (reaction no. 90). As a result, the excitation of O_2 molecules to the $b^1\Sigma_g^+$ state leads to an increase in the temperature at the end of the irradiation zone. The temperature behind the front also increases. For example, for $I_0 = 0$, the temperature is 1093 K, while for $I_0 = 10$ and 20 kW/cm^2, $T_1 = 1102$ and 1110 K, respectively. At the same time, for the same values of I_0, the value of T_1 in the case of excitation of state $a^1\Delta_g$ by radiation with $\lambda_I = 1.268$ μm does not exceed 1093 K.

The same situation is typical of a CH_4/O_2 mixture. However, the final temperature of the combustion products in this case is higher than in the CH_4/air mixture due to the fact that the absolute amount of methane in the stoichiometric mixture $CH_4/O_2 = 0.5/1$ is substantially larger than in the CH_4/air mixture ($CH_4/O_2/N_2 = 0.5/1/3.76$). On account of the larger concentration of O_2 molecules, the concentration of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules in the irradiation zone of the CH_4/O_2 mixture is higher than in the CH_4/air mixture with the same values of E_{in} and E_s. Consequently, in spite of the fact that the value of L_{in} in the absence of irradiation ($E_{in} = 0$) is virtually the same for the CH_4/O_2

and CH_4/air mixtures for the given parameters in the case of irradiation at $\lambda_I = 1.268$ μm, as well as at $\lambda_I = 762$ nm, the contraction of the induction zone in the CH_4/O_2 mixture is noticeably stronger than in the CH_4/air mixture, although the values of T_1 for the CH_4/O_2 mixture are even slightly lower (in the CH_4/O_2 mixture, $T_1 = 993$ and 1014 K for $I_0 = 10$ and 20 kW/cm^2, respectively). This is illustrated in Fig. 3 where the temperature variation in the irradiation zone ($\delta = 0.5$ m) and behind the shock wave front ($\beta = 30°$) is shown (Fig. 3a) for the mixture $CH_4/O_2 = 0.5/1$ and (Fig. 3b) for the mixture $CH_4/O_2/N_2 = 0.5/1/3.76$ with $T_0 = 300$ K, $P_0 = 10^4$ Pa, and $M_0 = 8$ under the action of radiation with $\lambda_I = 1.268$ μm and 762 nm for $I_0 = 10$ and 20 kW/cm^2 ($E_{in} \approx 2$ and 4 J/cm^2). It can be seen that the values of E_{in} required for the ignition of the CH_4/air mixture at admissible distances from the front (~1–3 m) are considerably higher (by a factor of 2–3) than for the CH_4/O_2 mixture. It is important to note that the excitation of O_2 molecules to the states $a^1\Delta_g$ and $b^1\Sigma_g^+$ does not lead to an increase in concentration of nitrogen oxides (NO and NO_2) in the combustion products.

The main reason for the contraction of the induction zone is naturally not the increase in T_1 under the action of radiation, but the intensification of the chain mechanism of combustion of the CH_4/O_2 mixture. It should be recalled that the main reaction of initiation of the chain at low temperatures ($T < 1300$ K) in CH_4/O_2 (air) mixtures is the interaction of CH_4 with O_2, during which an active radical CH_3 and hydrogen dioxide are formed. These compounds also interact with O_2 in the reactions $CH_3 + O_2(X^3\Sigma_g^-) = CH_2O + OH$, $CH_3 + O_2(X^3\Sigma_g^-) = CH_3O + O$ and $HO_2 + O_2(X^3\Sigma_g^-) = O_3 + OH$. The reaction products are chemically active radicals CH_2O and OH, as well as O atoms and O_3 (which gives atomic oxygen after dissociation). Oxygen atoms and OH radicals interact with CH_4, again forming CH_3 and OH ($CH_4 + O = CH_3 + OH$, $CH_4 + OH = CH_3 + H_2O$). However, for $T < 1100$ K, CH_3 radicals intensely recombine, forming passive compounds C_2H_6 and CH_3O_2 ($2CH_3 + M = C_2H_6 + M$; $CH_3 + O_2 + M = CH_3O_2 + M$). These reactions occur with energy release and lead to an increase in T. The rates of these reactions decrease and the concentration of CH_3 in the mixture increases. The rate of the reaction $CH_3 + O_2(X^3\Sigma_g^-) = CH_3O + O$ in which active O atoms are formed also increases. The decomposition of CH_3O radicals leads to the formation of CH_2O and H; i.e., this is a chain branching reaction. Hydrogen atoms reacting with O_2 yield O and OH ($H + O_2(X^3\Sigma_g^-) = OH + O$). The high concentration of active centers O, H, OH, and CH_3, as well as the formation of C_2H_5 radicals as a result of decomposition of passive radical C_2H_6, leads to the ignition of the mixture. However, in the absence of irradiation at $T \approx 1000$ K, all these processes occur at a very low rate and, hence, the length of the induction zone exceeds 200 m for $M_0 = 8$, $\beta = 30°$, and $P_0 = 10^4$ Pa.

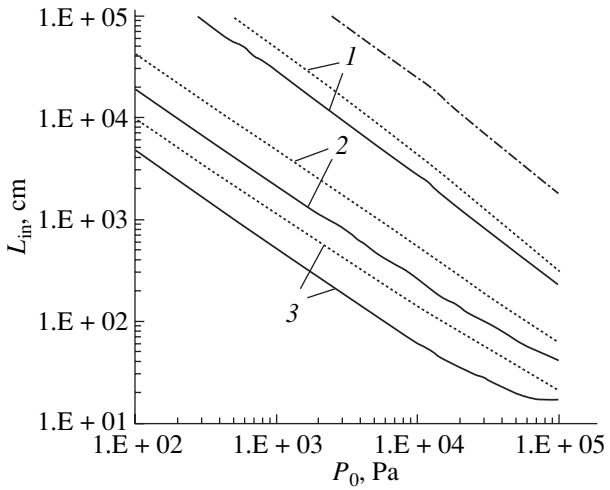


Fig. 4. Dependence of the induction zone length L_{in} on the initial pressure of the stoichiometric CH_4/air mixture for $T_0 = 300$ K, $M_0 = 8$, and $\beta = 30°$ upon the excitation of O_2 molecules by radiation with $\lambda_I = 1.268$ μm (solid curves) and 762 nm (dotted curves) for various values of the energy absorbed by an O_2 molecule: $E_s = 0.01$, 0.05, 0.1 eV/molecule (*1–3*). The dashed curve describes the dependence $L_{in}(P_0)$ in the absence of irradiation.

For excitation of O_2 molecules by radiation with $\lambda_I = 1.268$ μm as well as with $\lambda_I = 762$ nm, the scheme of the processes of formation of active atoms and radicals remains basically the same. However, the rates of all processes with the participation of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules increase by three to four orders of magnitude. This primarily refers to the chain initiation and continuation reactions involving $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules. Since the concentrations of all active radicals increase, the rates of even those reactions in which excited O_2 molecules are not involved directly also become higher. The contraction of the induction zone in this case is determined by the concentration of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules at the end of the irradiation zone (and, hence, on the value of E_{in}). The extent of this contraction also depends on the parameters P_0 and T_0 of the mixture in front of the shock wave; along with the value of E_{in}, these parameters determine the value of radiant energy E_s absorbed by an O_2 molecule. Figure 4 shows the dependence of L_{in} on P_0 for $M_0 = 8$, $\beta = 30°$, and $T_0 = 300$ K for various values of E_s under the action of radiation with $\lambda_I = 1.268$ μm and 762 nm. As expected, for the same value

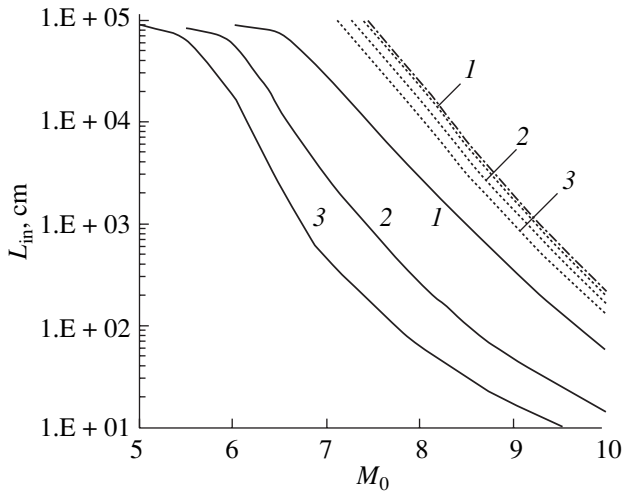


Fig. 5. Dependence of the induction length L_{in} on the Mach number for an unperturbed flow of the stoichiometric CH_4/air mixture with $T_0 = 300$ K, $P_0 = 10^4$ Pa, and $\beta = 30°$ in the case of excitation of O_2 molecules to the state $a^1\Delta_g$ by radiation with $\lambda_I = 1.268$ μm (solid curves) and in the case when the entire absorbed energy is spent on gas heating (dotted curves) for $E_s = 0.01$, 0.05, and 0.1 eV/molecule (*1–3*). The dot-and-dash curve corresponds to the absence of irradiation ($E_s = 0$).

of E_s, the contraction of the induction zone in the case of radiation with $\lambda_I = 1.268$ μm is stronger than for radiation with $\lambda_I = 762$ nm. Since the absorption length for radiation with $\lambda_I = 1.268$ μm is much larger than for radiation with $\lambda_I = 762$ nm (the values of L_v are equal to 6×10^4 and 610 cm, respectively, for $T_0 = 300$ K and $P_0 = 10^4$ Pa), a larger number of passes of a laser beam across the flow is required for radiation with $\lambda_I = 1.268$ μm to attain the same value of E_s. It is important for practical applications that, even for a high pressure of the gas mixture ($P_0 = 10^5$ Pa), the length L_{in} is reduced by a factor of 28 for a low value of the energy $E_s = 0.05$ eV/molecule absorbed by a single O_2 molecule (for $M_0 = 8$ and $\delta = 0.5$ m, this value of E_s corresponds to the laser radiation flux $E_{in} = 38$ J/cm^2 for radiation with $\lambda_I = 762$ nm), while the value of L_{in} itself for radiation with $\lambda_I = 762$ nm is equal to 0.6 m.

The method based on the excitation of O_2 molecules to the state $a^1\Delta_g$ or $b^1\Sigma_g^+$ by laser radiation is much more effective for initiating combustion and detonation in a supersonic flow than the method considered in [13, 19, 20], which is based on simple heating of the medium by laser radiation. Figure 5 shows the dependence $L_{in}(M_0)$ for various values of E_s under the action of radiation with $\lambda_I = 1.268$ μm on a CH_4/air mixture, in which O_2 molecules are excited to the state $a^1\Delta_g$ and the entire absorbed energy is spent on heating of the gas. It can be seen that, in the entire range of variation of M_0, the value of L_{in} in the case of excitation of O_2 molecules is one to two orders of magnitude smaller than in the case of purely thermal action of laser radiation even for a small value of $E_s = 0.05$ eV/molecule. To attain the same value of the length of the induction zone as in the case of excitation of O_2 molecules for $E_s = 0.1$ eV/molecule at $M_0 = 7$, 8 and $P_0 = 10^4$ Pa ($L_{in} = 1$ m), the energy that must be supplied to the gas under purely thermal action must be ~10 eV/molecule.

CONCLUSIONS

The action of laser radiation with a wavelength of $\lambda_I = 1.268$ μm or 762 nm on a supersonic flow makes it possible to initiate detonation mode of combustion in a CH_4/air mixture behind the front of relatively weak shock waves, when the gas temperature behind the front does not exceed 1100 K even for a low value of the absorbed energy ($E_s = 0.05$–0.1 eV/molecule). This is due to intensification of formation of active atoms O and H and radicals OH, CH_3, and CH_2O, which are carriers of the chain mechanism of combustion of methane–oxygen mixtures, due to the formation of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ excited molecules in the irradiation zone. In CH_4/air mixtures even at high values of pressure in front of the shock wave ($P_0 = 10^5$ Pa), the induction zone in the case of excitation of O_2 molecules to the states $a^1\Delta_g$ and to the $b^1\Sigma_g^+$ contracts by a factor of 25–30 for a low value of the absorbed energy ($E_s = 0.05$ eV/molecule) as compared to the case without irradiation. This makes it possible to realize detonation combustion in a supersonic flow at small distances (<1 m) from the irradiation zone in the CH_4/air mixture under the atmospheric pressure and at moderate values of temperature behind the shock wave ($T_1 \le 1100$ K). This method for initiating combustion in the supersonic flow is much more effective (by hundreds of time) than the thermal heating of the medium by resonant laser radiation.

ACKNOWLEDGMENTS

This study was financed by the Russian Foundation for Basic Research (project nos. 02-01-00703 and 02-02-16915).

REFERENCES

1. T. Yip, AIAA Pap., No. 89-2567 (1989).
2. N. G. Dautov and A. M. Starik, Fiz. Goreniya Vzryva **32**, 94 (1996).
3. C. Li, K. Kailasanath, and E. S. Oran, Combust. Flame **108**, 173 (1997).
4. L. Bezgin, A. Ganzhelo, O. Gouskov, and V. Kopchenov, AIAA Pap., No. 98-1513 (1998).

5. L. F. Figueira Da Silva and B. Deshaies, Combust. Flame **121**, 152 (2000).
6. D. E. Gonzalez, AIAA Pap., No. 96-4560 (1996).
7. W. Chinitz, AIAA Pap., No. 96-4536 (1996).
8. L. Bezgin, A. Ganzhelo, O. Gouskov, *et al.*, in *Gaseous and Heterogeneous Detonations. Science to Applications*, Ed. by G. Roy, S. Frolov, K. Kailasanath, and N. Smirnov (Enas, Moscow, 1999), pp. 285–300.
9. S. Eidelman, AIAA Pap., No. 99-1067 (1999).
10. A. M. Starik and N. S. Titova, Zh. Tekh. Fiz. **71** (8), 1 (2001) [Tech. Phys. **46**, 929 (2001)].
11. A. M. Starik and N. S. Titova, Dokl. Akad. Nauk **380**, 332 (2001) [Dokl. Phys. **46**, 627 (2001)].
12. A. M. Starik and N. S. Titova, in *Confined Detonations and Pulse Detonation Engines*, Ed. by G. Roy *et al.* (Torus, Moscow, 2003), pp. 87–104.
13. M. S. Chou, F. E. Fendell, and H. W. Behrens, Proc. SPIE **1862**, 45 (1993).
14. A. M. Starik and N. S. Titova, Fiz. Goreniya Vzryva **36** (3), 31 (2000).
15. T. M. Sloane, Combust. Sci. Technol. **63**, 287 (1989).
16. K. M. Leung and R. P. Lindstedt, Combust. Flame **102**, 129 (1995).
17. N. G. Dautov and A. M. Starik, Kinet. Katal. **38**, 207 (1997).
18. A. I. Didyukov, Yu. A. Kulagin, L. A. Shelepin, and V. N. Yarygina, Kvantovaya Élektron. (Moscow) **16**, 892 (1989).
19. M. J. McEwan and L. F. Phillips, *The Chemistry of the Atmosphere* (Halsted Press, New York, 1975; Mir, Moscow, 1978).
20. M. A. Tanoff, M. D. Smooke, R. E. Teets, and J. A. Fell, Combust. Flame **103**, 253 (1995).

Translated by N. Wadhwa

Technical Physics, Vol. 49, No. 9, 2004, pp. 1126–1134. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 23–31.

GASES AND LIQUIDS

Nonlinear Oscillations and Stability of a Charged Drop Moving Relative to a Dielectric Medium

V. A. Koromyslov, A. I. Grigor'ev, and S. O. Shiryaeva

Yaroslavl State University, Yaroslavl, 150000 Russia

e-mail: grig@uniyar.ac.ru

Received February 17, 2004

Abstract—Nonlinear calculations to within the second order of smallness with respect to the initial deformation of a liquid drop show that a stream of an ideal incompressible dielectric liquid streamlining the charged ideally conducting drop causes interaction between modes both in the first and second orders of smallness. Both the linear and nonlinear interactions of the oscillation modes result in the excitation of modes absent in the spectrum of the initial drop deformation. The relative motion of the drop and the medium leads to broadening of the spectrum of modes excited in the second order of smallness. The presence of the flow streamlining the drop and the intermode interaction result in decreasing the critical magnitudes of the drop charge and the velocity and density of the medium determining drop instability development. © *2004 MAIK "Nauka/Interperiodica".*

1. Charged drops moving relative to external media are encountered in various problems of technical physics and technology [1, 2]. In particular, this system is of interest in connection with investigations of the physical mechanism of lighting initiation [3, 4]. According to existing qualitative concepts, initiation of streak lighting is associated with ignition of a corona discharge in the vicinity of a large drop or water-bearing hailstone (accompanied by development of instability of the charged drop surface). However, such concepts have not found any substantiation in natural measurements in storm clouds, where maximum measured drop charges and internal electric fields [5] are much smaller than the necessary values predicted for the development of instability of the drop surface with respect to the initial and induced charges [6]. Apparently, in constructing the physical model of lighting initiation, a certain important factor (for example, the aerodynamic pressure in the vicinity of a falling drop) is omitted, which can account for reduction of the critical conditions of development of instability of a charged surface [7–9].

There are many papers devoted to analysis of the mechanism of disintegration of drops free falling through the atmosphere (see, for example, review [2] and references therein) that are mainly of experimental character. The rigorous analytical solution of the problem of oscillations and stability of a charged drop moving relative to the medium was found only in the linear approximation with respect to the magnitude of deformation of a spherical drop [7–9].

Analytical investigations of the nonlinear oscillations and stability of charged drops in a dielectric medium have been performed only for a drop immobile relative to the medium [10–12]. Attempts at nonlinear analysis of the stability of the surface of a charged drop moving relative to the medium are lacking so far. The present study attempts to fill this gap.

2. Let an ideal incompressible dielectric medium with a density ρ_2 and permittivity ε_* occupying infinite volume move at a constant velocity $\mathbf{U}_0$ relative to an immobile drop of an ideal incompressible ideally conducting liquid with a radius R. Denote the surface tension coefficient of the interface by σ, and the total charge of the drop by Q. Let us assume that, at the initial time instant $t = 0$, the equilibrium spherical shape of the drop undergoes a virtual axisymmetric perturbation of a finite amplitude which is significantly smaller than the drop radius. The perturbation is assumed to be proportional to one of the modes of capillary oscillations in the system. Our objective is to study nonlinear oscillations of the drop at $t > 0$.

To simplify the following calculations we shall employ dimensionless variables chosen in such a way that $R = \sigma = \rho_1 = 1$. Then, in a spherical coordinate system with the origin at the center of mass of the drop, the equation of the interface disturbed by axisymmetric wave motion will take the form

$$r = 1 + \xi(\theta, t), \quad |\xi| \ll 1.$$

Liquid motion in the drop and medium is assumed to be potential. That is, the velocity fields of the wave motions in the drop $\mathbf{V} = \nabla\psi(\mathbf{r}, t)$ and in the medium $\mathbf{U} = \nabla\varphi(\mathbf{r}, t)$ are defined by the functions of velocity potentials in the drop $\psi(\mathbf{r}, t)$ and in the medium $\varphi(\mathbf{r}, t)$.

The mathematical statement of the problem of calculation of the nonlinear oscillations of the drop–medium interface includes:

1063-7842/04/4909-1126$26.00 © 2004 MAIK "Nauka/Interperiodica"

Laplace equations for the velocity potentials $\psi(\mathbf{r}, t)$ and $\varphi(\mathbf{r}, t)$, and electrostatic potential $\Phi(\mathbf{r}, t)$

$$\Delta\Phi(\mathbf{r}, t) = 0; \quad \Delta\psi(\mathbf{r}, t) = 0; \quad \Delta\varphi(\mathbf{r}, t) = 0 \quad (1)$$

and the corresponding boundary conditions at the drop center

$$r \longrightarrow 0: \quad \psi(\mathbf{r}, t) \longrightarrow 0; \quad (2)$$

at infinity

$$r \longrightarrow \infty: \quad \Phi(\mathbf{r}, t) \longrightarrow 0; \quad \nabla\varphi(\mathbf{r}, t) \longrightarrow \mathbf{U}_0; \quad (3)$$

and at the interface:

$$r = 1 + \xi: \quad \frac{\partial\xi}{\partial t} = \frac{\partial\psi}{\partial r} - \frac{1}{r^2}\frac{\partial\psi}{\partial\theta}\frac{\partial\xi}{\partial\theta}, \quad (4)$$

including kinematic condition of equality of the normal liquid velocity components in the drop and medium

$$\frac{\partial\psi}{\partial r} - \frac{1}{r^2}\frac{\partial\psi}{\partial\theta}\frac{\partial\xi}{\partial\theta} = \frac{\partial\varphi}{\partial r} - \frac{1}{r^2}\frac{\partial\varphi}{\partial\theta}\frac{\partial\xi}{\partial\theta} \quad (5)$$

and dynamic condition of constancy of the electric potential at the drop surface

$$-\frac{\partial\psi}{\partial t} - \frac{1}{2}(\nabla\psi)^2 + P_{\text{in}} + P_E - P_\sigma$$

$$= -\rho\frac{\partial\varphi}{\partial t} - \frac{\rho}{2}(\nabla\varphi)^2 + P_{\text{ex}}; \quad (6)$$

$$P_E = \frac{\varepsilon_*(\nabla\Phi)^2}{8\pi}; \quad P_\sigma = \text{div}\mathbf{n};$$

$$\Phi(\mathbf{r}, t) = \Phi_S(t). \quad (7)$$

In these relationships, P_{in} and P_{ex} are the pressures inside the drop and in the medium, respectively, P_E is the pressure of the electric field of the drop charge at the interface, P_σ is the Laplace pressure, n is the unit vector of a positive normal to the drop surface, $\Phi_S(t)$ is the electric potential constant at the drop surface, and $\rho_2/\rho_1 \equiv \rho$.

In addition to the boundary conditions listed above, it is also necessary to take into account the conditions of constancy of the drop charge

$$-\frac{\varepsilon_*}{4\pi}\oint_S(\mathbf{n}\cdot\nabla\Phi)dS = Q; \quad S = \begin{cases} r = 1 + \xi(\theta, t) \\ 0 \le \theta \le \pi \\ 0 \le \vartheta \le 2\pi \end{cases} \quad (8)$$

and constancy of the drop volume

$$\int_{V_1} r^2 dr\sin\theta d\theta d\vartheta = \frac{4}{3}\pi; \quad V_1 = \begin{cases} 0 \le r \le 1 + \xi(\theta, t) \\ 0 \le \theta \le \pi \\ 0 \le \vartheta \ge 2\pi. \end{cases} \quad (9)$$

The condition of immobility of the center of mass of the system, as shown in [12], for sufficiently large linear scales of the medium is fulfilled automatically; therefore, calculation of the amplitude of the transnational (fundamental) mode, as well as higher modes, should be performed on the basis of hydrodynamic boundary conditions at the interface.

We formulate the initial conditions assuming axisymmetric deformation of a spherical drop and zero initial velocity of the interface motion at the initial instant:

$$\xi(\theta, t = 0) = \xi_0 P_0(\cos\theta) + \varepsilon P_k(\cos\theta) \ (k \ge 2); \quad (10)$$

$$\frac{\partial\xi(\theta, t = 0)}{\partial t} = 0. \quad (11)$$

Here ε is the amplitude of the initial perturbation, which is a small parameter of the problem; $P_k(\cos\theta)$ is the Legendre polynomial of the kth order; and ξ_0 is a constant defined by condition (9) which in the second order of smallness takes the form

$$\xi_0(\theta, t = 0) = -\varepsilon^2\frac{1}{(2k + 1)} + O(\varepsilon^3). \quad (12)$$

3. To solve problem (1)–(12) in the quadratic approximation with respect to the oscillation amplitude, we employ a multiscale method. For this purpose, we represent the unknown functions $\xi(\theta, t)$, $\psi(\mathbf{r}, t)$, $\varphi(\mathbf{r}, t)$, $\Phi(\mathbf{r}, t)$ in the form of series in powers of the small parameter ε and series in Legendre polynomials similar to those described in [10, 11]. In doing so we assume that these functions are dependent not on the time t as such, but on different time scales $T_m = \varepsilon^m t$

$$\xi(\theta, t) = \sum_{m=1}^{\infty}\varepsilon^m\xi^{(m)}(\theta, T_0, T_1, \ldots);$$

$$\psi(\mathbf{r}, t) = \sum_{m=0}^{\infty 1}\varepsilon^m\psi^{(m)}(r, \theta, T_0, T_1, \ldots);$$

$$\varphi(\mathbf{r}, t) = \sum_{m=0}^{\infty}\varepsilon^m\varphi^{(m)}(r, \theta, T_0, T_1, \ldots); \quad (13)$$

$$\Phi(\mathbf{r}, t) = \sum_{m=0}^{\infty}\varepsilon^m\Phi^{(m)}(r, \theta, T_0, T_1, \ldots).$$

We shall calculate derivatives with respect to time keeping in mind the complete set of different scales according to the rule

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial T_0} + \varepsilon\frac{\partial}{\partial T_1} + O(\varepsilon^2).$$

Substituting expansions (13) into boundary-value problem (1)–(9) and equating terms of the same order of smallness in each of the equations, one can obtain a set of boundary-value problems for successive determi-

nation of the functions $\xi^{(m)}$, $\varphi^{(m)}$, $\psi^{(m)}$, $\Phi^{(m)}$, where $m = 0, 1, 2, \ldots$.

4. In the zeroth order of smallness, problem (1)–(9) takes the form

$$\Delta\Phi^{(0)} = 0; \quad \Delta\varphi^{(0)} = 0;$$
$$r \longrightarrow \infty: \Phi^{(0)} \longrightarrow 0; \quad \nabla\varphi^{(0)} \longrightarrow \mathbf{U}_0;$$
$$r = 1: \frac{\partial\varphi^{(0)}}{\partial r} = 0; \tag{14}$$
$$P_{\text{in}} + \frac{\varepsilon_*}{8\pi}\left(\frac{d\Phi^{(0)}}{dr}\right)^2 - 2 = -\frac{\rho}{2}\left(\frac{\partial\varphi^{(0)}}{\partial\theta}\right)^2 + P_{\text{ex}};$$
$$\Phi^{(0)} = \Phi_S^{(0)}; \quad -\frac{\varepsilon_*}{4\pi}\int_0^{2\pi}\int_0^{\pi}\frac{d\Phi^{(0)}}{dr}\sin\theta d\theta d\vartheta = Q.$$

Solutions (14) describing the equilibrium state of the system have the form

$$\varphi^{(0)}(\theta) \equiv U\left(r + \frac{1}{2r^2}\right)\cos\theta;$$
$$\Phi_S^{(0)} \equiv \frac{Q}{\varepsilon_*}; \quad \Phi^{(0)} \equiv \frac{Q}{\varepsilon_* r}. \tag{15}$$

5. Due to linearity of equations (1)–(3), each of functions $\varphi^{(m)}$, $\psi^{(m)}$, and $\Phi^{(m)}$ in expansions (13) must satisfy these equations; therefore, we represent these functions for $m \geq 1$ in the form of series in Legendre polynomials

$$\psi^{(m)}(r, \theta, T_0, T_1, \ldots) = \sum_{n=0}^{\infty} E_n^{(m)}(T_0, T_1, \ldots) r^n P_n(\cos\theta);$$
$$\varphi^{(m)}(r, \theta, T_0, T_1, \ldots) = \sum_{n=0}^{\infty} G_n^{(m)}(T_0, T_1, \ldots) r^{-n-1} P_n(\cos\theta); \tag{16}$$
$$\Phi^{(m)}(r, \theta, T_0, T_1, \ldots) = \sum_{n=0}^{\infty} F_n^{(m)}(T_0, T_1, \ldots) r^{-n-1} P_n(\cos\theta).$$

In similar form we shall search for successive corrections $\xi^{(m)}$ to the expression defining the drop surface shape

$$\xi^{(m)}(\theta, T_0, T_1, \ldots) = \sum_{n=0}^{\infty} M_n^{(m)}(T_0, T_1, \ldots) P_n(\cos\theta). \tag{17}$$

To determine unknown coefficients $G_n^{(1)}$, $E_n^{(1)}$, $F_n^{(1)}$, $M_n^{(1)}$ in solutions (16), (17) (for $m = 1$) in the first order of smallness with respect to ε, the set of boundary and initial conditions (4)–(11) subject to (12) should be transformed to

$$r = 1: \frac{\partial\xi^{(1)}}{\partial T_0} = \frac{\partial\psi^{(1)}}{\partial r};$$
$$\frac{\partial\psi^{(1)}}{\partial r} = \frac{\partial\varphi^{(1)}}{\partial r} + \xi^{(1)}\frac{\partial^2\varphi^{(0)}}{\partial r^2} - \frac{\partial\varphi^{(0)}}{\partial\theta}\frac{\partial\xi^{(1)}}{\partial\theta};$$
$$-\frac{\partial\psi^{(1)}}{\partial T_0} + \frac{\varepsilon_*}{4\pi}\frac{d\Phi^{(0)}}{dr}\left(\frac{d\Phi^{(1)}}{dr} + \xi^{(1)}\frac{d^2\Phi^{(0)}}{dr^2}\right) + (2 + \Delta_\Omega)\xi^{(1)}$$
$$= \rho\left\{-\frac{\partial\varphi^{(1)}}{\partial T_0} + \xi^{(1)}\left(\frac{d\varphi^{(0)}}{\partial\theta}\right)^2 - \frac{\partial\varphi^{(0)}}{\partial\theta}\frac{\partial\varphi^{(1)}}{\partial\theta}\right\}; \tag{18}$$
$$\Phi^{(1)} + \frac{d\Phi^{(0)}}{dr}\xi^{(1)} = \Phi_S^{(1)};$$
$$\int_0^{\pi}\left[\frac{\partial\Phi^{(1)}}{dr} + \left(\frac{d^2\Phi^{(0)}}{dr^2} + 2\frac{d\Phi^{(0)}}{dr}\right)\xi^{(1)}\right]\sin\theta d\theta = 0;$$
$$t = 0: \ \xi^{(1)} = P_k(\cos\theta); \ \frac{\partial\xi^{(1)}}{\partial T_0} = 0; \ \int_0^{\pi}\xi^{(1)}\sin\theta d\theta = 0.$$

Here, Δ_Ω is the angular part of the Laplace operator. Substituting expansions (16) and (17) (at $m = 1$) as well as solutions (15) of the zeroth order of smallness into set (18) makes it possible to obtain an infinite set of coupled differential equations defining unknown coefficients $G_n^{(1)}(T_0, T_1)$, $E_n^{(1)}(T_0, T_1)$, $F_n^{(1)}(T_0, T_1)$, $M_n^{(1)}(T_0, T_1)$

$$M_0^{(1)}(T_0, T_1, \ldots) \equiv 0; \quad M_1^{(1)}(T_0, T_1, \ldots) \equiv 0;$$
$$n \geq 2; \quad A_n M_{n-2}^{(1)}(T_0, T_1, \ldots) + B_n\frac{\partial M_{n-1}^{(1)}(T_0, T_1, \ldots)}{\partial T_0}$$
$$+ \frac{\partial^2 M_n^{(1)}(T_0, T_1, \ldots)}{\partial T_0^2} + \omega_n^2 M_n^{(1)}(T_0, T_1, \ldots)$$
$$+ C_n\frac{\partial M_{n+1}^{(1)}(T_0, T_1, \ldots)}{\partial T_0} + D_n M_{n+2}^{(1)}(T_0, T_1, \ldots) = 0;$$
$$E_n^{(1)}(T_0, T_1, \ldots) = \frac{1}{n}\frac{\partial M_n^{(1)}(T_0, T_1)}{\partial T_0};$$
$$G_n^{(1)}(T_0, T_1, \ldots) = -\frac{1}{n+1}\frac{\partial M_n^{(1)}(T_0, T_1)}{\partial T_0}$$

$$+\frac{2}{3}U\left[\frac{n}{(2n-1)}M_{n-1}^{(1)}(T_0, T_1) - \frac{n}{(2n+3)}M_{n+1}^{(1)}(T_0, T_1)\right]; \quad (19)$$

$$A_n = \frac{9}{4}U^2\rho\chi(n)\frac{n^2(n-1)(n-2)}{(2n-3)(2n-1)};$$

$$B_n = \frac{3}{2}U\rho n\chi(n); \quad C_n = \frac{3}{2}Y\rho\chi(n)\frac{n(2n+1)}{n2+3};$$

$$D_n^{(1)} = \frac{9}{4}U^2\rho\chi(n)\frac{n^2(n+1)(n+2)}{(2n+3)(2n+5)};$$

$$\omega_n^2 = \chi(n)\left(n(n-1)(n+2-W)\right.$$

$$\left.-\frac{9}{2}U^2\rho\frac{n^2((2n+1)(n^2-1)+3)}{(2n-1)(2n+1)(2n+3)}\right);$$

$$W \equiv \frac{Q^2}{4\pi\varepsilon_*}; \quad \chi(n) = \left(1+\rho\frac{n}{n+1}\right)^{-1};$$

$$F_n^{(1)}(T_0, T_1, \ldots) = \frac{Q}{\varepsilon_*}M_n^{(1)}(T_0, T_1, \ldots); \quad \Phi_S^1 \equiv 0.$$

Set (19) allows one to determine the dependence of the coefficients of the expansions only on time scale T_0. Their dependences on other time scales are determined in the next orders of smallness.

It is easy to see that at $U = 0$; that is, in the case of an immobile medium, the linear interaction of modes determined by Eqs. (19) vanishes. The set of coupled differential equations (19) decomposes into a set of independent differential equations of the second order with constant coefficients defining harmonic oscillations of separate modes (as was obtained earlier [10, 11] for the oscillations of a charged drop of an incompressible liquid at rest relative to an ideal incompressible dielectric medium). Thus, the reason for the emergence of the linear interaction of modes with respect to the small parameter is the motion of the medium. According to (19), the *n*th mode interacts with four nearest modes, namely, with $(n-2)$th, $(n-1)$th, $(n+1)$th, and $(n+2)$th ones. Previously, the interaction between modes in the linear approximation with respect to a small parameter was found in the case of a plane interface between immiscible ideal incompressible media, one of which performed translational motion parallel to the interface [13], that is, in a situation when the interface is capable of featuring the Kelvin–Helmholtz instability. In [7–9], it was shown that in a case of a drop streamlined by the ideal liquid, the drop surface is involved in the oscillatory motion typical of this instability.

Note also one more effect of the interaction between a drop and the flow of streamlining ideal liquid, found in the linear approximation. According to [14, 15] the drop is flattened along the stream into a spheroid with an eccentricity depending on the flow velocity and the drop charge. Possible oscillations of the drop will occur in the vicinity of the equilibrium spheroidal shape. However, the degree of sphericity at reasonable velocities (for which the flow of the medium streamlining the drop can be assumed to be laminar) is usually very small. According to [14, 15], the amplitude of the spheroidal deformation is given by expression $M_2^{(1)} = (3\rho_2 RU^2/16\sigma)$, and, for example, in calculations of air flow around a drop with $R = 100$ μm when $\rho_2 \approx 0.001$ g/cm for any reasonable flow velocities (until $\mathrm{Re} = (RU/\nu) \le 20$, where ν is the kinematic air viscosity) the spheroidal deformation can be neglected in the second order of smallness.

To complete the problem consideration in the linear approximation with respect to ε, it is possible to assume the values of $M_n^{(1)}(T_0, T_1)$ to be independent of the time scale T_1, that is, to represent them in the form $M_n^{(1)}(T_0, T_1) \approx M_n^{(1)}(T_0) + O(T_1)$. In doing so, we obtain the following estimate for the surface perturbation:

$$\xi(\theta, t) = \varepsilon\xi^{(1)}(\theta, t) + \varepsilon O(\varepsilon t). \quad (20)$$

This expansion is uniformly applicable at $t \le O(\varepsilon^{-1})$. For $t \ge O(\varepsilon^{-1})$, the given expansion becomes inapplicable. Thus, expression (20) is valid on a time interval $t \le O(1)$, and, in this case, the error is on the order of $\sim\varepsilon^2$. However, when studying the tendencies of the surface motion, one can use expression (20) also on a time interval $t \le O(\varepsilon^{-1})$, provided that the solution of the first order is comparable with the magnitude of the initial deformation.

6. In order to determine corrections of the second order of smallness, that is, to find the functions $G_n^{(2)}$, $E_n^{(2)}$, $F_n^{(2)}$, and $M_n^{(2)}$ from the set of boundary conditions (4)–(12), we keep in (13) the terms of the second order of smallness in ε and obtain the following equations

$$r = 1: \frac{\partial\xi^{(2)}}{\partial T_0} + \frac{\partial\xi^{(1)}}{\partial T_1} = \frac{\partial\psi^{(2)}}{\partial r} + \xi^{(1)}\frac{\partial^2\psi^{(1)}}{\partial r^2} - \frac{\partial\xi^{(1)}}{\partial\theta}\frac{\partial\psi^{(1)}}{\partial\theta};$$

$$\frac{\partial\psi^{(2)}}{\partial r} + \xi^{(1)}\frac{\partial^2\psi^{(1)}}{\partial r^2} - \frac{\partial\xi^{(1)}}{\partial\theta}\frac{\partial\psi^{(1)}}{\partial\theta} = \frac{\partial\varphi^{(2)}}{\partial r} + \xi^{(1)}\frac{\partial^2\varphi^{(1)}}{\partial r^2}$$

$$-\frac{\partial\xi^{(1)}}{\partial\theta}\frac{\partial\varphi^{(1)}}{\partial\theta} + \xi^{(2)}\frac{\partial^2\varphi^{(0)}}{\partial r^2} + \frac{1}{2}(\xi^{(1)})^2\frac{\partial^3\varphi^{(0)}}{\partial r^3}$$

$$-\frac{\partial\xi^{(2)}}{\partial\theta}\frac{\partial\varphi^{(0)}}{\partial\theta} + 2\xi^{(1)}\frac{\partial\xi^{(1)}}{\partial\theta}\frac{\partial\varphi^{(0)}}{\partial\theta};$$

$$-\frac{\partial\psi^{(2)}}{\partial T_0}-\frac{\partial\psi^{(0)}}{\partial T_1}-\xi^{(1)}\frac{\partial^2\psi^{(1)}}{\partial r\partial T_0}-\frac{1}{2}\left[\left(\frac{\partial\psi^{(1)}}{\partial r}\right)^2+\left(\frac{\partial\psi^{(1)}}{\partial\theta}\right)^2\right]$$
$$+\frac{\varepsilon_*}{8\pi}\left[2\frac{\partial\Phi^{(0)}}{\partial r}\frac{\partial\Phi^{(2)}}{\partial r}+\left(\frac{\partial\Phi^{(1)}}{\partial r}\right)^2+\left(\frac{\partial\Phi^{(1)}}{\partial\theta}\right)^2\right.$$
$$+2\xi^{(2)}\frac{\partial\Phi^{(0)}}{\partial r}\frac{\partial^2\Phi^{(0)}}{\partial r^2}+2\xi^{(1)}\left(\frac{\partial\Phi^{(0)}}{\partial r}\frac{\partial^2\Phi^{(1)}}{\partial r^2}+\frac{\partial^2\Phi^{(0)}}{\partial r^2}\frac{\partial\Phi^{(1)}}{\partial r}\right)$$
$$\left.+(\xi^{(1)})^2\left(\frac{\partial\Phi^{(0)}}{\partial r}\frac{\partial^3\Phi^{(1)}}{\partial r^3}+\left(\frac{\partial^2\Phi^{(0)}}{\partial r^2}\right)^2\right)\right]+[(2+\Delta_\Omega)\xi^{(2)}$$
$$-2\xi^{(1)}(1+\Delta_\Omega)\xi^{(1)}]=\rho\left\{-\frac{\partial\varphi^{(2)}}{\partial T_0}-\frac{\partial\varphi^{(1)}}{\partial T_1}-\xi^{(1)}\frac{\partial^2\varphi^{(1)}}{\partial r\partial T_0}\right.$$
$$-\frac{1}{2}\left[\left(\frac{\partial\varphi^{(1)}}{\partial r}\right)^2+\left(\frac{\partial\varphi^{(1)}}{\partial\theta}\right)^2\right]-\frac{1}{2}(\xi^{(1)})^2\left(\frac{\partial^2\varphi^{(0)}}{\partial r^2}\right)^2$$
$$-\frac{(\xi^{(1)})^2}{2}\left[3\left(\frac{\partial\varphi^{(0)}}{\partial\theta}\right)^2+\frac{\partial\varphi^{(0)}}{\partial\theta}\frac{\partial^3\varphi^{(0)}}{\partial r^2\partial\theta}\right]-\frac{\partial\varphi^{(0)}}{\partial\theta}\frac{\partial\varphi^{(2)}}{\partial\theta}$$
$$\left.+\xi^{(1)}\left[2\frac{\partial\varphi^{(0)}}{\partial\theta}\frac{\partial\varphi^{(1)}}{\partial\theta}-\frac{\partial\varphi^{(0)}}{\partial\theta}\frac{\partial^2\varphi^{(1)}}{\partial r\partial\theta}\right]\right\};$$
$$\Phi^{(2)}+\xi^{(2)}\frac{d\Phi^{(0)}}{dr}+\xi^{(1)}\frac{\partial\Phi^{(1)}}{\partial r}+\frac{1}{2}(\xi^{(1)})^2\frac{d^2\Phi^{(0)}}{dr^2}=\Phi_S^{(2)};$$
$$\int_0^\pi\left[\frac{\partial\Phi^{(2)}}{\partial r}+\xi^{(1)}\left(\frac{\partial^2\Phi^{(1)}}{\partial r^2}+2\frac{\partial\Phi^{(1)}}{\partial r}\right)\right.$$
$$+\xi^{(2)}\left(\frac{d^2\Phi^{(0)}}{dr^2}+2\frac{d\Phi^{(0)}}{dr}\right)+(\xi^{(1)})^2\left(\frac{1}{2}\frac{d^3\Phi^{(0)}}{dr^3}\right.$$
$$\left.\left.+2\frac{d^2\Phi^{(0)}}{dr^2}+\frac{d\Phi^{(0)}}{dr}\right)-\frac{\partial\Phi^{(1)}}{\partial\theta}\frac{\partial\xi^{(1)}}{dr}\right]\sin\theta d\theta=0;$$
$$t=0:\ \xi_0^{(2)}=-\frac{1}{(2k+1)};\quad\frac{\partial\xi^{(2)}}{\partial T_0}+\frac{\partial\xi^{(1)}}{\partial T_1}=0;$$
$$\int_0^\pi[\xi^{(2)}+(\xi^{(1)})^2]\sin\theta d\theta=0.$$

Substituting expansions (16), (17) (at $m=2$), and solutions (15) and (19) into (21), we write a set of inhomogeneous differential equations for determination of unknown coefficients $G_n^{(2)}$, $E_n^{(2)}$, $F_n^{(2)}$, and $M_n^{(2)}$. Eliminating all terms proportional to $(\partial M_m^{(1)}(T_0, T_1, \ldots)/\partial T_1)$ that cause emergence of secular terms in the solutions, we find that amplitudes $M_m^{(1)}$ of the expansion are independent of the time scale T_1. Thus, in the following we can assume that $M_m^{(1)}(T_0, T_1, \ldots)\approx M_m^{(1)}(T_0)+O(T_2)$. Based on the above considerations, the set of differential equations for the coefficients $G_n^{(2)}$, $E_n^{(2)}$, $F_n^{(2)}$, and $M_n^{(2)}$ is reduced to the following:

$$M_0^{(2)}(T_0, T_1, \ldots)=\sum_{n=2}^{\infty}\frac{1}{2n+1}(M_0^{(1)}(T_0))^2;$$
$$A_nM_{n-2}^{(2)}(T_0, T_1, \ldots)+B_n\frac{\partial M_{n-1}^2(T_0, T_1, \ldots)}{\partial T_0}$$
$$+\frac{\partial^2M_n^{(2)(T_0, T_1, \ldots)}}{\partial T_0^2}+\omega_n^2M_n^{(2)}(T_0, T_1, \ldots)$$
$$+C_n\frac{\partial M_{n+1}^{(2)}(T_0, T_1, \ldots)}{\partial T_0}$$
$$+D_nM_{n+2}^{(2)}(T_0, T_1, \ldots)=\chi(n)f_n(T_0);\quad n\geq 1;$$
$$E_n^{(2)}(T_0, T_1, \ldots)=\frac{1}{n}\left(\frac{\partial M_n^{(2)}(T_0, T_1, \ldots)}{\partial T_0}\right.$$
$$\left.-\sum_{n=2}^{\infty}\sum_{l=2}^{\infty}\left[(m-1)K_{m,l,n}-\frac{\alpha_{m,l,n}}{m}\right]\frac{\partial M_m^{(1)}(T_0)}{\partial T_0}M_l^{(1)}(T_0)\right);$$
$$G_n^{(2)}(T_0, T_1, \ldots)=\frac{3n}{2(2n-1)}UM_{n-1}^{(2)}(T_0, T_1, \ldots)$$
$$-\frac{3n}{2(2n+3)}UM_{n+1}^{(2)}(T_0, T_1, \ldots)-\frac{1}{n+1}\frac{\partial M_n^{(2)}(T_0, T_1, \ldots)}{\partial T_0}$$
$$-\sum_{m=2}^{\infty}\sum_{l=2}^{\infty}\left\{\left[\frac{m+2}{n+1}K_{m,l,n}-\frac{\alpha_{m,l,n}}{(n+1)(m+1)}\right]\right.$$
$$\times\frac{\partial M_m^{(1)}(T_0)}{\partial T_0}M_l^{(1)}(T_0)+\frac{3}{2}UM_m^{(1)}(T_0)M_l^{(1)}(T_0)$$
$$\times\left[-\frac{m(m-1)^2}{(n+1)(2m+1)}K_{m-1,l,n}+\frac{(m+1)^2(m+2)}{(n+1)(2m+1)}K_{m+1,l,n}\right.$$
$$\left.\left.-\frac{(m+1)\alpha_{m+1,l,n}}{(n+1)(2m+1)}+\frac{(m-1)\alpha_{m-1,l,n}}{(n+1)(2m+1)}\right]\right\};$$
$$F_n^{(2)}(T_0, T_1, \ldots)=\frac{Q}{\varepsilon_*}\left\{M_n^{(2)}(T_0, T_1, \ldots)\right.$$
$$\left.+\sum_{m=2}^{\infty}\sum_{l=2}^{\infty}(mK_{m,l,n}M_m^{(1)}(T_0)M_l^{(1)}(T_0))\right\};$$
$$\Phi_S^{(2)}\equiv 0;$$

$$f_n(T_0) = \sum_{m=2}^{\infty}\sum_{l=2}^{\infty}\{M_m^{(1)}(T_0)M_l^{(1)}(T_0)$$
$$\times\Bigg\{K_{m,l,n}\Bigg(2n[l(l+1)-1]$$
$$+W\frac{n}{2}[l(m+1)-m(2m-2n+7)+3]$$
$$+U^2\rho\frac{n(9m^2+9m-7)}{(2m-1)(2m+3)}\Bigg)+\alpha_{m,l,n}W\frac{n}{2}$$
$$+U^2\rho\frac{9n(n-1)}{4(2n-1)(2m+1)}[m(m-1)^2K_{m-1,l,n-1}$$
$$-(m+1)^2(m+2)K_{m+1,l,n-1}-(m-1)\alpha_{m-1,l,n-1}$$
$$+(m+1)\alpha_{m+1,l,n-1}]+U^2\rho\frac{9n(n+1)}{4(2n+3)(2m+1)}$$
$$\times[(m+1)^2(m+2)K_{m+1,l,n+1}-m(m-1)^2K_{m-1,l,n+1}$$
$$-(m+1)\alpha_{m+1,l,n+1}+(m-1)\alpha_{m-1,l,n+1}] \quad (22)$$
$$-U^2\rho\frac{9n}{8(2m+1)}\Bigg[\frac{m(m-1)}{2m+1}K_{m-2,l,n}$$
$$+\frac{(m+1)(m+2)}{2m+3}K_{m+2,l,n}\Bigg]\Bigg\}+U^2\rho\frac{9nm(m+1)}{4(2m+1)}$$
$$\times\left(\frac{M_{m-1}^{(1)}(T_0)M_l^{(1)}(T_0)}{2m-1}-\frac{M_{m+1}^{(1)}(T_0)M_l^{(1)}(T_0)}{2m+3}\right)$$
$$\times[(m-1)(m+2)K_{m+1,l,n}-m(m+5)K_{m-1,l,n}]$$
$$+\frac{9}{8}U^2\rho n[(m+1)(l+1)K_{m,l,n}+\alpha_{m,l,n}]$$
$$\times\left(\frac{mM_{m-1}^{(1)}(T_0)}{2m-1}-\frac{mM_{m+1}^{(1)}(T_0)}{2m+3}\right)\left(\frac{lM_{l-1}^{(1)}(T_0)}{2l-1}-\frac{lM_{l+1}^{(1)}(T_0)}{2l+3}\right)$$
$$+\frac{3}{4}U\rho n[(m+1)(l+1)K_{m,l,n}+\alpha_{m,l,n}]\Bigg[\frac{m}{l+1}\Bigg(\frac{M_{m+1}^{(1)}(T_0)}{2m+3}$$
$$\times\frac{\partial M_l^{(1)}(T_0)}{T_0}-\frac{M_{m-1}^{(1)}(T_0)}{2m-1}\frac{\partial M_l^{(1)}(T_0)}{\partial T_0}\Bigg)+\frac{l}{m+1}$$
$$\times\left(\frac{\partial M_m^{(1)}(T_0)}{\partial T_0}\frac{M_{l+1}^{(1)}(T_0)}{2l+3}-\frac{\partial M_m^{(1)}(T_0)}{\partial T_0}\frac{M_{l-1}^{(1)}(T_0)}{2l-1}\right)\Bigg]$$
$$+\Bigg[(m-n-1)K_{m,l,n}-\frac{\alpha_{m,l,n}}{m}-\frac{\rho n}{n+1}\Bigg((m-n+1)K_{m,l,n}$$
$$-\frac{\alpha_{m,l,n}}{m+1}\Bigg)\Bigg]\frac{\partial M_m^{(1)}(T_0)}{\partial T_0}M_l^{(1)}(T_0)+\Bigg[\left(m-1-\frac{n}{2}\right)K_{m,l,n}$$
$$-\frac{n+2l}{2ml}\alpha_{m,l,n}+\frac{\rho n}{2(n+1)}\Bigg((n-2m-3)K_{m,l,n}$$
$$+\frac{(n+2l+3)\alpha_{m,l,n}}{(m+1)(l+1)}\Bigg)\Bigg]\frac{\partial M_m^{(1)}(T_0)}{\partial T_0}\frac{\partial M_l^{(1)}(T_0)}{\partial T_0}+\frac{3}{2}U\rho n$$
$$\times\frac{\partial M_m^{(1)}(T_0)}{\partial T_0}M_l^{(1)}(T_0)\Bigg(\frac{m(4+5n+3m+mn+m^2)}{(n+1)(2m+1)}K_{m-1,l,n}$$
$$+\frac{(m+2)(2+n+m-mn+m^2)}{(n+1)(2m+1)}K_{m+1,l,n}$$
$$+\frac{(n-1)(m+2)}{2n-1}K_{m,l,n-1}-\frac{(n+1)(m+2)}{2n+3}K_{m,l,n+1}$$
$$+\frac{(m-1)\alpha_{m-1,l,n}}{(n+1)(2m+1)}-\frac{(n-1)\alpha_{m,l,n-1}}{(m+1)(2n-1)}+\frac{(n+1)\alpha_{m,l,n+1}}{(m+1)(2n+3)}$$
$$-\frac{(n+1)\alpha_{m+1,l,n}}{(n+1)(2m+1)}\Bigg)+U\rho\frac{3n}{2(n+1)(2m+1)}$$
$$\times[(m+1)^2(m+2)K_{m+1,l,n}-m(m-1)^2K_{m-1,l,n}$$
$$-(m+1)\alpha_{m+1,l,n}+(m-1)\alpha_{m-1,l,n}]M_m^{(1)}(T_0)\frac{\partial M_l^{(1)}(T_0)}{\partial T_0}$$
$$+\frac{3}{2}mU\rho m(m+1)K_{m,l,n}\Bigg(\frac{\partial M_{m+1}^{(1)}(T_0)}{\partial T_0}\frac{M_l^{(1)}(T_0)}{2m+3}$$
$$-\frac{\partial M_{m-1}^{(1)}(T_0)}{\partial T_0}\frac{M_l^{(1)}(T_0)}{2m-1}\Bigg)\Bigg\};$$
$$K_{mln}\equiv[C_{m0l0}^{n0}]^2;$$
$$\alpha_{mln}\equiv-\sqrt{m(m+1)l(l+1)}C_{m0l0}^{n0}C_{m-1l0}^{n0};$$

where C_{m0l0}^{n0} and C_{m-1l0}^{n0} are the Clebsch–Gordan coefficients.

Consideration of the problem in the quadratic approximation with respect to ε allows one to determine the dependence of coefficients $M_n^{(2)}(T_0, T_1)$ only on the time scale T_0. This makes it possible to write $M_n^{(2)}(T_0, T_1) \approx M_n^{(2)}(T_0) + O(T_1)$ and obtain the following estimate for the surface perturbation:

$$\xi(\theta, t) = \varepsilon\xi^{(1)}(\theta, t) + \varepsilon^2\xi^{(2)}(\theta, t) + O(\varepsilon^3 t). \quad (23)$$

Expression (23) is valid on a time interval $t \le O(1)$ to an accuracy of $\sim\varepsilon^3$. On a time interval $O(1) \le t \le O(\varepsilon^{-1})$, the error becomes comparable to the second term (with correction of the second order of smallness); consequently, in expansion (23) only the first term remains valid that corresponds to the linear approximation. Thus, the approximate solution of the linear problem (20) is applicable (uniformly valid) on a time interval $t \le O(\varepsilon^{-1})$.

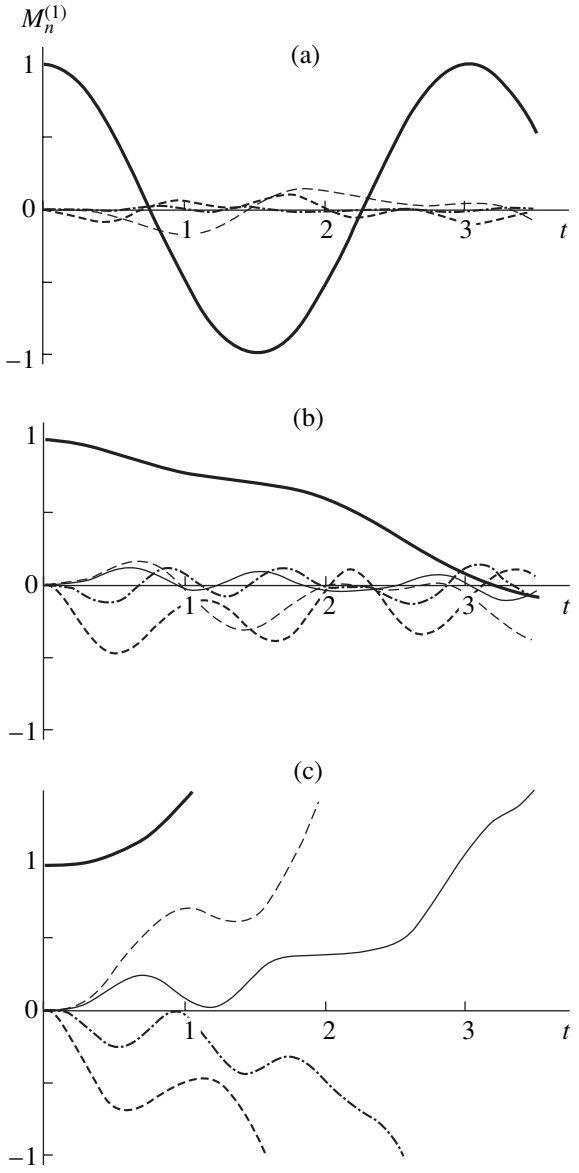


Fig. 1. The dimensionless amplitude coefficients $M_n^{(1)}(t)$ of modes excited in the first order of smallness as functions of dimensionless time for the initial drop deformation determined by the fundamental mode ($k = 2$) when $W = 1$ and $\rho = 0.1$. Thick solid curve refers to the second mode; thin dashed line, to the third mode; dashed line of medium thickness, to the fourth mode; dot-and-dash line, to the fifth mode; and thin solid line, to the sixth mode. $U = 2$ (a), 4 (b), and 4.5 (c).

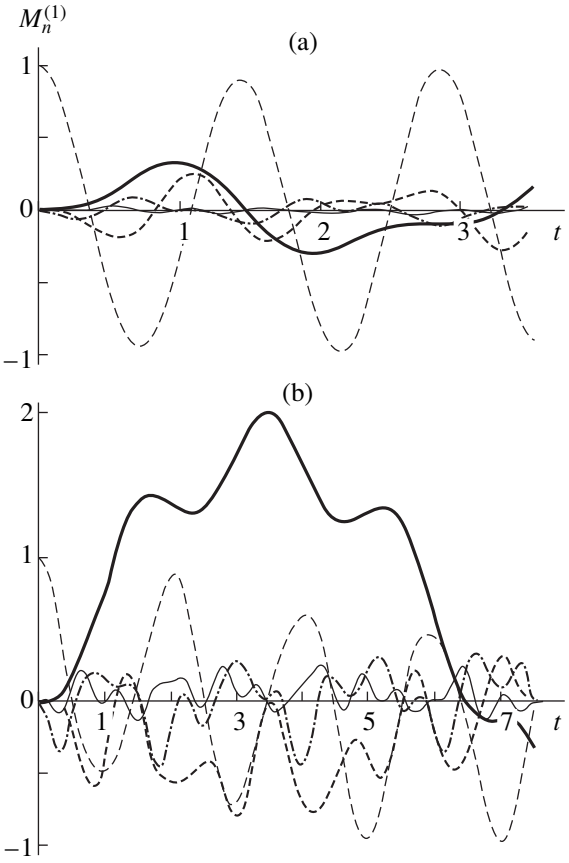


Fig. 2. Same as in Fig. 1 for the initial deformation determined by the third mode ($k = 3$). $U = 2$ (a) and 4 (b).

7. Numerical solution of the set of differential equations (19) for $M_n^{(1)}(T_0)$ obtained with the aid of the program package Mathematika (restricting the consideration to the first five modes, $n = 2, 3, 4, 5, 6$) is illustrated in Figs. 1 and 2. These dependences show that, at low velocities U of the medium in the sense of the drop stability with respect to the intrinsic charge, a noticeable contribution to the spectrum of the capillary oscillations is made only by the initially excited mode ($n = k$), and (at $k \neq 2$) the fundamental mode (which is excited automatically due to interaction with the flow as a result of redistribution of the hydrodynamic pressure at the drop surface [14, 15]). Note that the velocity values critical for the development of instability are determined by condition $\omega_2^2 = 0$

$$\left(n(n-1)(n+2-W)\right.$$

$$\left. -\frac{9}{2}U^2\rho\frac{n^2((2n+1)(n^2-1)+3)}{(2n-1)(2n+1)(2n+3)}\right) = 0 \quad \text{at} \quad n = 2.$$

The contribution of the other modes ($n \neq k$) determined by the linear intermode interaction, according to [19], is small. In this case, the drop surface performs oscillations, which are close to harmonic and correspond to superposition of the kth (initially excited) mode and the fundamental mode in the vicinity of the equilibrium shape. If the velocity of the medium is close to the critical value for which the drop becomes unstable ($\omega_2^2 < 0$), then the contribution of the other modes (excited due to the linear interaction) to the drop shape becomes more noticeable (as seen from the com-

parison between Figs. 1a, 2a and 1b, 2b. In this case, modes that are closer in number n to the initially excited kth mode possess greater oscillation amplitudes, which decrease with increasing n.

Note that, in the dimensionless variables employed, the velocity is normalized to $\sqrt{(\sigma/(R\rho_1))}$; that is, a dimensionless velocity of $U = 1$ corresponds to the dimensional velocity $U = \sqrt{(\sigma/(R\rho_1))}$. For example, at $U = 1$, for a drop with radius $R = 100$ μm in an air flow, the dimensional velocity is ≈84 cm/s. The Reynolds number Re = (UR/ν) for such a drop is ≈5, and air flow in the vicinity of the drop is laminar.

Figures 1a–1c and 2a, 2b present the results of calculations with the use of Eqs. (19), showing the time variation of the amplitudes of modes excited in the first order of smallness due to the intermode interaction when the initial deformation is determined by a virtual excitation of the fundamental ($k = 2$) mode (Fig. 1) and the third ($k = 3$) mode (Fig. 2). The calculations were performed for various values of the velocity U of the medium.

According to Fig. 1, the oscillation amplitudes of the modes excited due to the linear interaction grow with increasing velocity. In Fig. 1c, it is seen that, when the second mode is initially unstable, some nearest modes coupled to it by the linear interaction according to (19) become unstable due to the intermode interaction, in spite of the fact that these modes (at the magnitudes of U, W, ρ used in the calculations) must retain stability in the sense of the linear theory because $\omega^2_{n\neq 2} > 0$. It should be noted that the velocity used in the calculations (Fig. 1c) only slightly exceeds the critical value for which the drop becomes unstable (at $W = 1$ and ρ = 0.1, this critical value for the fundamental mode is $U_{cr} = 4.4$). It is interesting to note that, with further growth of the flow velocity (at $U = 5$), the magnitude of the instability increment of the third mode exceeds that of the fundamental mode, and the third mode starts to play the main part in the development of instability [7].

From the above analysis, it follows that, under otherwise equal conditions and the initial excitation of a higher (not fundamental) mode, the critical velocity of the medium from the standpoint of instability development exhibits a decrease. For example, calculations show that, for the initial excitation of the third mode, the fundamental mode becomes unstable already at $U = 4.05$. In other words, the presence of the intermode interaction results in decreasing magnitudes of parameters U, W, and ρ for which the drop becomes unstable.

Figure 2 demonstrates the results of calculating the amplitude coefficients of modes excited due to the linear interaction in the case when the initial deformation is determined by the third mode. From a comparison of Figs. 2a, 2b, 2c with Figs. 1a, 1b it is seen that the amplitude of the fundamental mode (determined under these conditions both by the action of the hydrodynamic pressure and by the interaction with initially excited third mode) may be comparable with the amplitude of the initially excited mode (Fig. 2a) and may even exceed the latter (Fig. 2b).

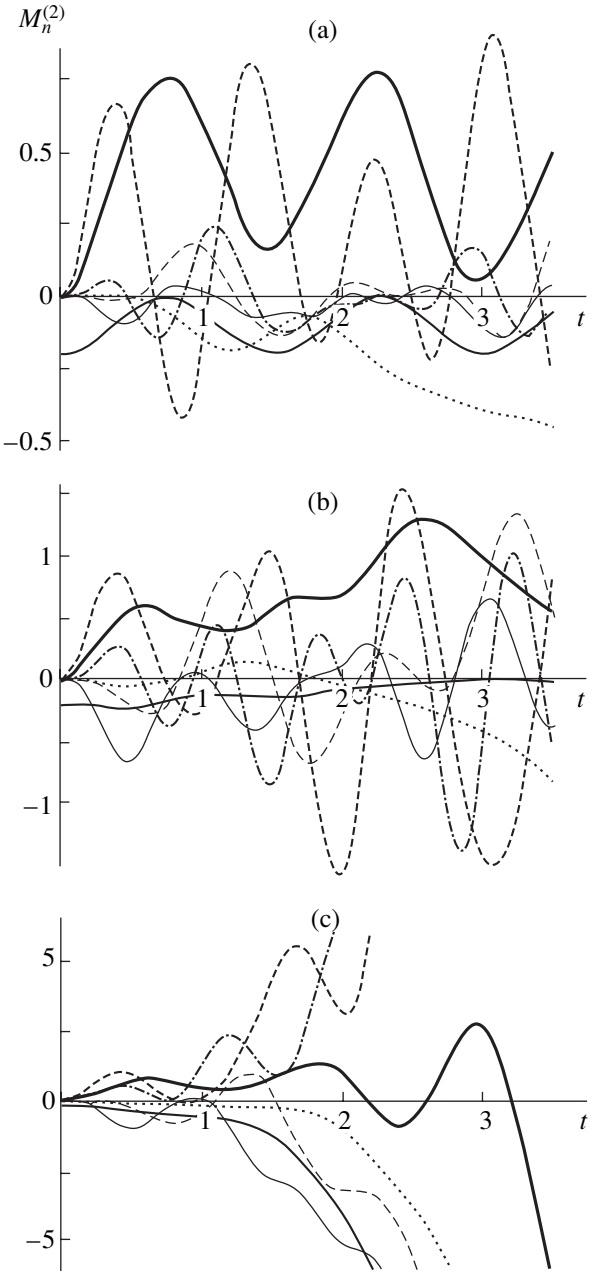


Fig. 3. The dimensionless amplitude coefficients $M_n^{(2)}(t)$ of modes excited due to the nonlinear interaction in the second order of smallness for the initial drop deformation determined by the principal mode ($k = 2$) when $W = 1$ and ρ = 0.1. The solid curve of a medium thickness refers to the zeroth mode; thin short, dashed line, to the first mode; the other lines correspond to the legends in Fig. 2. $U = 2$ (a), 4 (b), and 4.5 (c).

Numerical solution of the set of differential equations (19) and (22) for the amplitudes $M_n^{(2)}(T_0)$ of modes excited in the second order of approximations due to the nonlinear interaction (for the modes from zeroth to sixth) shows that, at low velocities U of the medium, by analogy with the linear approximation, the largest amplitudes are observed for the modes that would be excited for the immobile medium [9–11], that is, for modes with the numbers $n = 2j$, where $j = 0, 1, \ldots, k$. The motion of the medium results in exciting additional modes in the second order of smallness whose emergence is associated with the presence of modes in the spectrum of the first order of smallness, which are different from the initially excited mode and appear only due to the linear interaction. The amplitudes of such additionally excited modes are very small and their contribution to the shape of an oscillating drop is insignificant.

Figures 3a–3c show the time variation of modes excited in the second order of smallness due to the nonlinear interaction when the initial deformation is determined by a virtual perturbation of the fundamental ($k = 2$) mode. The calculations were performed using Eqs. (22) for various velocities U of the medium. When the medium is immobile, only the zeroth, second, and fourth modes would be excited in the second order of smallness. The presence of motion of the medium results in additional excitation of the first, third, fifth, and sixth modes. Nevertheless, it is seen that the amplitudes of the fundamental and forth modes are maximum.

In Fig. 3, it is also seen that the amplitudes of the initially excited modes increase with the velocity of the medium. Figures 1c and 3c demonstrate that, along with the linear instability of modes, the modes excited in the second order of smallness also become unstable, and the instability of modes that are higher than the fundamental one exhibit oscillatory character. It is interesting to note that, at $n \geq 3$, the amplitudes of the identical modes in the first and second approximations have opposite signs.

In the numerical analysis of sets (19) and (22) the consideration was restricted only to the first several modes (corresponding to the first several equations in the infinite sets of equations), the amplitudes of which most significantly influence the shape of the drop surface, whereas the complete solution must include each infinite spectrum of modes associated with the interaction according to Eqs. (19) and (22). However, the amplitudes of higher modes are very small and rapidly decrease as their numbers become much greater than the number of the mode excited at the initial moment.

CONCLUSIONS

The presence of an ideal incompressible dielectric medium streamlining a charged ideally conducting drop results in the emergence of interaction between oscillation modes both in the first and second orders of smallness. This interaction leads to the excitation of modes that are absent in the spectrum of modes determining the initial deformation of the drop. As the flow velocity increases, the oscillation amplitudes of the initially excited modes grow. Relative motion of the drop and the medium, as well as the mode interaction result in decrease of the critical magnitudes of the intrinsic charge of the drop, as well as the velocity and density of the medium from the standpoint of drop instability development.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 03-01-00760.

REFERENCES

1. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
2. A. L. Gonor and V. Ya. Rivkind, Itogi Nauki Tekh., Ser.: Mekh. Zhidk. Gaza **17**, 98 (1982).
3. V. A. Dyachuk and V. M. Muchnik, Dokl. Akad. Nauk SSSR **248**, 60 (1979).
4. A. I. Grigor'ev and S. O. Shiryaeva, Phys. Scr. **54**, 660 (1996).
5. *Clouds and Cloud Atmosphere: A Handbook,* Ed. by I. P. Mazina, A. Kh. Khrgian, and I. M. Imyanitov (Gidrometeoizdat, Leningrad, 1989) [in Russian].
6. A. I. Grigor'ev, S. O. Shiryaeva, and S. I. Shchukin, Zh. Prikl. Khim. (St. Petersburg) **72**, 117 (1999).
7. A. I. Grigor'ev, V. A. Koromyslov, and S. O. Shiryaeva, Zh. Tekh. Fiz. **69** (5), 7 (1999) [Tech. Phys. **44**, 486 (1999)].
8. A. I. Grigor'ev, V. A. Koromyslov, and S. O. Shiryaeva, Zh. Tekh. Fiz. **70** (7), 26 (2000) [Tech. Phys. **45**, 840 (2000)].
9. V. A. Koromyslov and A. I. Grigor'ev, Zh. Tekh. Fiz. **72** (9), 21 (2002) [Tech. Phys. **47**, 1090 (2002)].
10. V. A. Koromyslov, S. O. Shiryaeva, and A. I. Grigor'ev, Zh. Tekh. Fiz. **73** (9), 44 (2003) [Tech. Phys. **48**, 1124 (2003)].
11. M. V. Rybakova, S. O. Shiryaeva, and A. I. Grigor'ev, Zh. Tekh. Fiz. **74** (1), 24 (2004) [Tech. Phys. **49**, 22 (2004)].
12. S. O. Shiryaeva, A. I. Grigor'ev, V. A. Koromyslov, and A. N. Zharov, Zh. Tekh. Fiz. **73** (9), 60 (2003) [Tech. Phys. **48**, 1141 (2003)].
13. S. O. Shiryaeva, Zh. Tekh. Fiz. **71** (3), 9 (2001) [Tech. Phys. **46**, 280 (2001)].
14. A. I. Grigor'ev, Zh. Tekh. Fiz. **72** (7), 41 (2002) [Tech. Phys. **47**, 834 (2002)].
15. A. I. Grigor'ev, V. A. Koromyslov, and M. V. Rybakova, Elektrokhim. Obrab. Met., No. 6, 22 (2002).

Translated by N. Mende

Technical Physics, Vol. 49, No. 9, 2004, pp. 1135–1141. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 32–38.

GAS DISCHARGES, PLASMA

Experimental Study of the Effect of a Longitudinal Magnetic Field on the Cathode Parts of a Glow Discharge in Helium

I. M. Ulanov and A. Yu. Litvintsev

Institute of Thermal Physics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 1, Novosibirsk, 630090, Russia

e-mail: litvintsev@itp.nsc.ru

Received June 30, 2003

Abstract—The effect of a longitudinal magnetic field on the electrical and optical properties of negative glow and the Faraday dark space of a glow discharge in helium is experimentally studied at a discharge current J_d = 11 mA and a pressure of 100–150 Pa in the discharge chamber. An experimental setup is designed and described. A longitudinal electric field is found to be absent in the region of negative glow and the Faraday dark space in both the presence and the absence of the magnetic field. The magnetic field is shown to decrease the voltage of the glow discharge and to increase the total discharge luminosity and the intensities of some spectral lines by a factor of 10–12. The experimental results are explained. *© 2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Although glow discharge has been studied in many works, the processes that occur in its cathode regions, which are located between the cathode and the positive column and include the layers of cathode drop, negative glow, and Faraday dark space, are still not completely clear [1–3].

The main feature of the longitudinal structure of a glow discharge is that the electron energy spectrum in the region of negative glow and Faraday dark space has a nonlocal character. As was shown in [4], three electron groups with mean energies of 2, 22.5, and 150 eV were detected in a helium discharge. In this case, simple probe measurements can lead to wrong results.

The most important point is that, in these regions of negative glow and Faraday dark space, an electric field is absent (or it is very small). Therefore, the regions of negative glow and Faraday dark space are most sensitive to an external action on a glow discharge [3]. We have a plasma free of an electric field; in this case, a magnetic field affects this plasma in the pure form.

Many experimental studies are dealing with the effect of a longitudinal magnetic field on a glow discharge. In most of such publications, the positive column of a discharge was studied in a longitudinal magnetic field [5–7]. In these works, however, the optical characteristics of the positive column of a glow discharge have not been studied.

The related investigations are described in [8, 9]. However, the authors of [8] did not indicate the place of measurements along the discharge, and Rokhlin [9] performed optical measurements in the region of the positive column of a discharge.

The purpose of this work is to study the effect of a longitudinal magnetic field on the optical characteristics (total emission and the emission of some spectral lines) of a glow discharge in the region of negative glow and Faraday dark space.

EXPERIMENTAL

For measurements, we designed an experimental setup shown schematically in Fig. 1. Discharge quartz tube *1* had an inner diameter of 45 mm, and the distance between the anode and cathode was 56 cm. Cathode *2* was made of molybdenum 42 mm in diameter, and the anode was made of a stainless steel 42 mm in diameter. The system was continuously evacuated by forepump *6* with a nitrogen trap to maintain a stable discharge pressure and to remove impurities that enter into the discharge from the quartz tube of discharge chamber *3* and the electrodes. The gas to be studied was supplied from

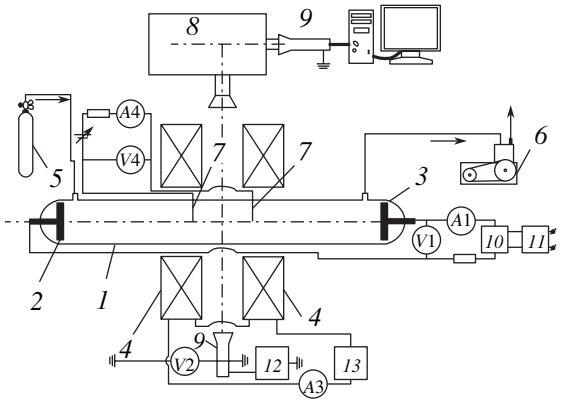


Fig. 1. Schematic diagram of the experimental setup: (*8*) monochromator, (*9*) photomultiplier, (*10*) IP-1, (*11*) RNO, (*12*) IP-2, and (*13*) IP-3.

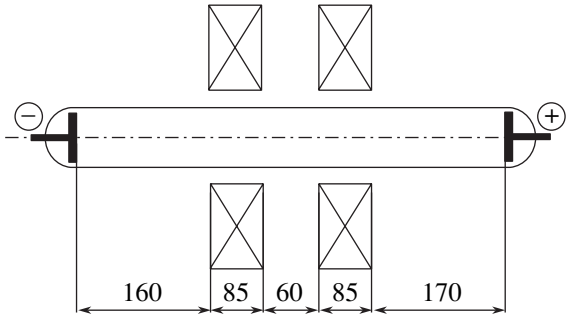


Fig. 2. Arrangement of the coils for generating a magnetic field.

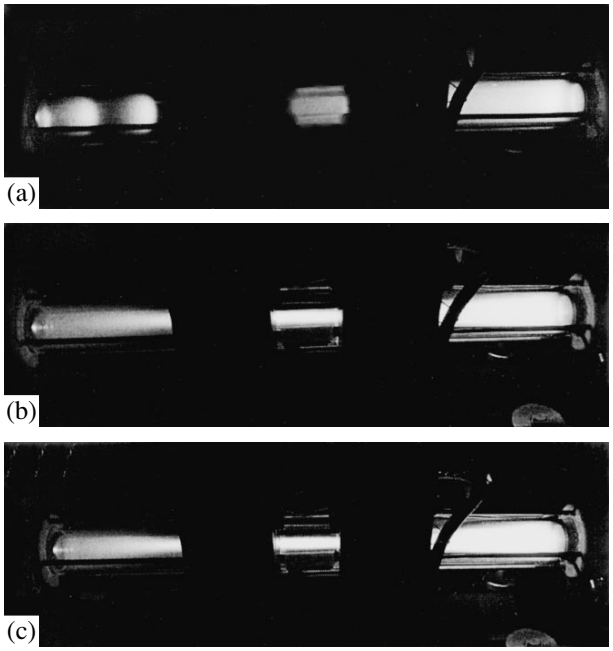


Fig. 3. Photographs of the discharge (a) without a magnetic field and in a magnetic field of (b) 0.025 and (c) 0.05 T.

vessel *5* through a needle valve. The pressure in the discharge chamber was measured with a VD-1 elastic element pressure gauge.

Magnetic system *4* consisted of two coils and generated a longitudinal magnetic field up to 0.1 T. To ensure the observation of a discharge and optical measurements, the coils had a 6-cm-wide gap. The magnetic field in the gap between the coils had a gradient of less than 5% of the field at the center of each coil. The magnetic system was graduated with an RSh1-10 magnetometer. The magnetic induction during experiments was maintained accurate to 1% by an IP-3 power supply with a controlled direct current.

The glow discharge was powered by an IP-1 power supply with a controlled voltage (up to 5 kV). The voltage and current were measured with an electrostatic voltmeter *V*1 and an ammeter *A*1 with an accuracy of 1 and 2%, respectively. The total discharge emission was measured with an FEU-2 device graduated using a reference TRU 1100-2350 thermal lamp. Spectral measurements were performed with an MRD-23 monochromator, and signals were applied to an analog-to-digital converter interfaced with a computer.

For probe measurements, we placed two probes *7*, which were silver wires 0.6 cm in diameter insulated by a thin quartz capillary, in the discharge chamber. The working surface of one probe was 2 mm^2. The probes were placed at the center of the discharge chamber. The interprobe distance was 8 cm; therefore, the temperature and electron density were averaged over this distance.

We only applied the two-probe method of measuring the temperature and electron density since it was the most reasonable method for measurements in a magnetic field.

EXPERIMENTAL RESULTS

To correctly estimate the results of visual observation of the effect of a magnetic field on the glow discharge and to interpret the photographs given below, we show the scheme of location of the coils generating the magnetic field with respect to the anode and cathode (Fig. 2).

In zero magnetic field, all characteristic parts of a glow discharge were clearly visible in the discharge tube. At the pressures and discharge currents under study, the positive column was always stratified, and, at a pressure of ~150 Pa, its length was about 20 cm. As the pressure increased, the length of the positive column increased and it could occupy almost the whole length of the discharge tube.

When even a very weak magnetic field (~0.005 T) was applied, the positive column of the discharge disappeared at a pressure less than 100–130 Pa in the discharge tube, and the cathode parts of the discharge, such as cathode glow, negative glow, and Faraday dark space, occupied the whole length of the tube. In this case, the discharge began to contract, and the diameter of the glowing part decreased with increasing magnetic field strength. This is clearly visible in Fig. 3, where photographs of the discharge without a magnetic field and with a magnetic field of 0.05 T at a pressure of 100 Pa in the discharge tube and a discharge current of 11 mA are given. The dark bands in the photographs are the tightening studs of the setup, which shade the discharge.

When studying the electrical characteristics of the discharge, we measured the total voltage *U* across the discharge, the voltage between the probe and cathode, and the electric field strength between the probes. Within the limits of experimental error, we failed to detect an electric field between the probes when they were in the region of negative glow and the Faraday dark space (at any rate, the field was less than 0.01 V/cm).

The results of measuring the total voltage across the discharge and the voltage between the probe and the anode are shown in Fig. 4. It is seen that, as the magnetic field increases, the total voltage across the discharge decreases and the voltage between the probe and the anode decreases sharply to a certain constant value. For a discharge current of 11 mA and a pressure of 130 Pa in the discharge chamber, this value is 5.4 V. This is likely to be the anode drop, and the initial sharp decrease in the voltage between the probe and the anode can be explained by the loss of the positive column, in which an electric field exists, in the discharge with increasing magnetic field.

It is interesting to note that, if the pressure in the discharge chamber increases to 200–250 Pa at the same discharge current (11 mA), the magnetic field does not affect the electrical characteristics of the discharge. The total voltage across the discharge and the voltage between the probe and the anode remain constant, the discharge does not contract, and the length of the positive column does not change.

Passing to the results of probe measurements, we have to note that we used only a double probe, since the application of a single probe for measurements in a magnetic field causes serious complications during interpretation of the results obtained.

Double probes can be used in rather strong magnetic fields [10, 11], which is due to the fact that an ion saturation current plays a decisive role for such a probe.

Figure 5 shows the typical probe characteristics for a double probe at a discharge current of 11 mA and a pressure of 130 Pa in the discharge chamber. The probe characteristics were always recorded so that the probe nearest to the anode was not in the positive column of the discharge.

The electron temperature was calculated by the formula [10]

$$\frac{kT_e}{e} = \frac{J_1^i J_2^i}{J_1^i + J_2^i}\left(\frac{dJ}{dU}\right)^{-1},$$

where J_1^i is the ion saturation current of the first probe, J_2^i is the ion saturation current of the second probe, and $(dJ/dU)^{-1}$ is the derivative of the total probe current.

The measured electron temperature versus the magnetic field strength is shown in Fig. 6. Since the probe characteristics shown in Fig. 5 have a symmetrical shape up to a certain value of the magnetic field, the electron temperature in the gap between the probes is assumed to be constant. This assumption is supported by visual observations: the discharge column compressed by the magnetic field glows uniformly over the entire gap between the probes. The probe characteristics at a magnetic field higher than 0.05 T become asymmetric, since the electron temperatures at each probe are different. In this case, the mean electron temperature in the gap between the probes is measured. The

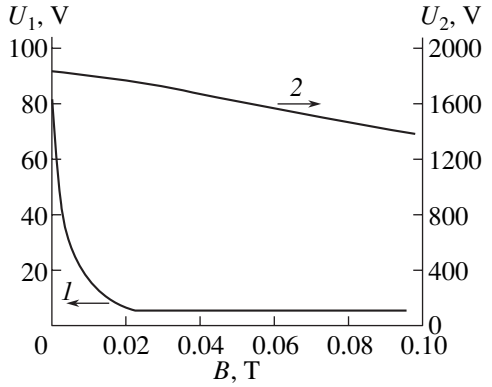


Fig. 4. Dependences of (*1*) the total voltage across the discharge, U_2, and (*2*) the voltage between the probe and the anode, U_1, on the magnetic field (J_d = 11 mA, P = 130 Pa).

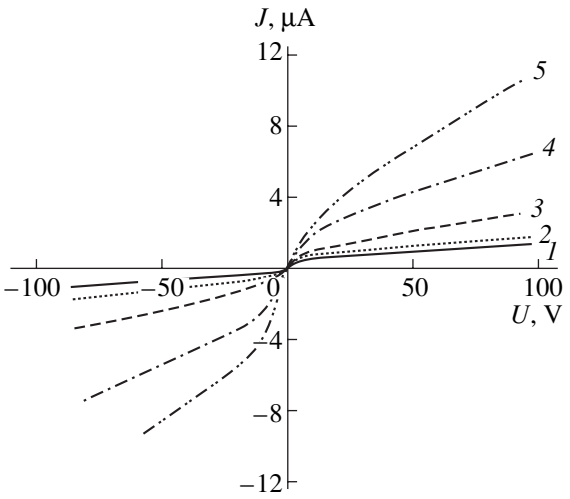


Fig. 5. Probe characteristics of the discharge in He at J_d = 11 mA and P = 130 Pa: (*1*) in zero magnetic field and in a magnetic field of (*2*) 0.005, (*3*) 0.01, (*4*) 0.025, and (*5*) 0.05 T.

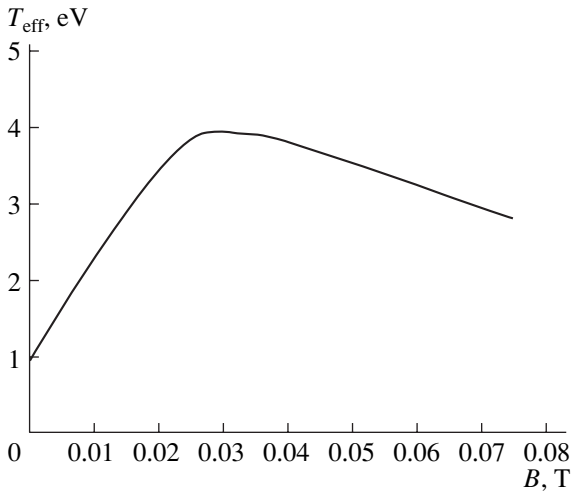


Fig. 6. Electron temperature T_{eff} as a function of the magnetic field (J_d = 11 mA, P = 130 Pa).

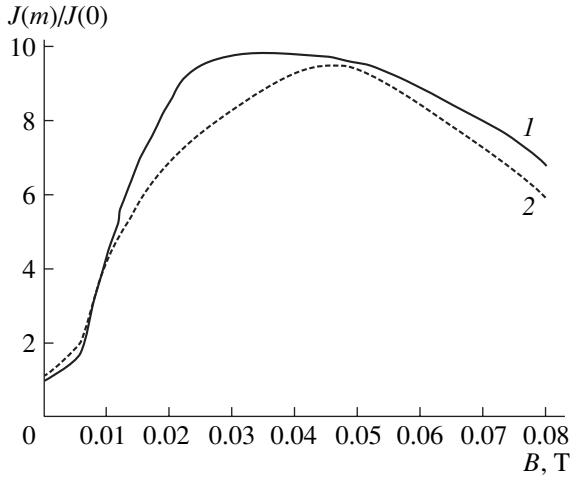


Fig. 7. Dependence of the relative intensity of the discharge glow on the magnetic field strength in two spectral ranges: (*1*) 320–1200 and (*2*) 340–620 nm.

luminosity of the discharge column near each probe is seen to be different. It is interesting to note that the authors of [5–7, 9] studied the effect of a magnetic field on the positive column of a discharge and found that the electron temperature decreased with increasing field. In [9], at the same magnetic field strength, the electron temperature decreased by more than half. This comparison again shows that a glow discharge is nonuniform along its length.

Let us now consider the results of optical measurements. Figure 7 shows the dependence of the relative intensity of a discharge glow on the magnetic induction. Here, $J(0)$ is the intensity of the discharge glow in zero magnetic field, and $J(M)$ is the same intensity but at various values of magnetic induction.

A spectral range of 340–1200 nm was chosen in order to exclude the long-wave radiation of the heated parts of the setup and the instruments used. To this end, we applied a water filter. The spectral range of visible radiation (340–620 nm) was selected by attaching a CZC23 glass filter to the water filter.

As is seen from Fig. 7, the total discharge luminosity increases strongly with the magnetic induction. The curves have a maximum at a magnetic field of 0.02–0.05 T. The same maximum is detected in the electron temperature in Fig. 5.

To find out how the magnetic field affects the intensity of some spectral lines, we studied the discharge spectra in the wavelength range 350–750 nm. Figure 8 shows the spectrum of He in zero magnetic field and in the presence of a 0.025-T magnetic field. It is seen that, as the magnetic field increases, the intensities of the spectral lines increase. Like the total intensity of a discharge glow (Fig. 7), the intensities of the spectral lines are maximal at 0.02–0.05 T.

The intensities of the He spectral lines in a longitudinal magnetic field increase differently with the wavelength: they increase by a factor of 10–12 in the short-wave range and by a factor of 8–10 in the long-wave range.

The lines of hydrogen appear in the He spectrum due to hydrogen emanation from the cathode. These impurity lines of hydrogen are clearly visible in Fig. 9. Taking into account the large difference in the excitation potentials of hydrogen and helium, it is obvious that even low contents of hydrogen impurities cause the emergence of hydrogen spectral lines. Figure 9 shows the helium spectrum at a pressure of 180 Pa in the discharge chamber. In this case, the positive column of the discharge is seen in the observation region between the magnetic coils. As follows from Fig. 9, mainly hydrogen impurity lines are excited. The effect of the magnetic field leads to a decrease in the line intensities. This result is consistent with the data reported in [9], where the positive column of a discharge was studied.

An interesting result was obtained when the He spectrum of the ac discharge was studied. Unfortunately, only a few experiments were performed with an alternating current. The discharge current and the pressure in the discharge chamber were maintained identical to those in the case of a direct current.

Figure 10 shows the spectrum of He discharge at an alternating current of 11 mA in magnetic fields of 0.025 and 0.05 T. Visual observation of the ac discharge demonstrates cathode glow at both electrodes, which transforms into negative glow. At the center of the discharge tube, only negative glow is observed. At a pressure of 100 Pa, there is no positive column. When a longitudinal magnetic field is applied, the discharge begins to contract from both electrodes. As the magnetic field increases, the diameter of the glow discharge filament decreases. As the pressure increases, a stratified positive column appears at the center of the tube. When the pressure is above 150 Pa, the length of the positive column increases, and, at a pressure of ~300 Pa, it occupies ~80% of the length of the discharge tube.

The behavior of the discharge spectrum is also unusual. As the magnetic field increases (Fig. 10), some spectral lines cease to be excited and some lines, on the contrary, are excited, and so on.

Based on the results obtained, we can assume that, at a certain pressure in the discharge chamber, a longitudinal magnetic field in the ac discharge exerts a sufficiently strong effect on the electron temperature and the energy distribution function of electrons.

DISCUSSION OF THE RESULTS

As noted above, the authors of [3–7, 9] studied the effect of a longitudinal magnetic field on the positive column. We found that the effect of a longitudinal magnetic field on the positive column of the discharge is very weak at a discharge current of 11 mA, a magnetic induction up to 0.05 T, and a helium pressure above 180 Pa in the discharge chamber. We did not detect a

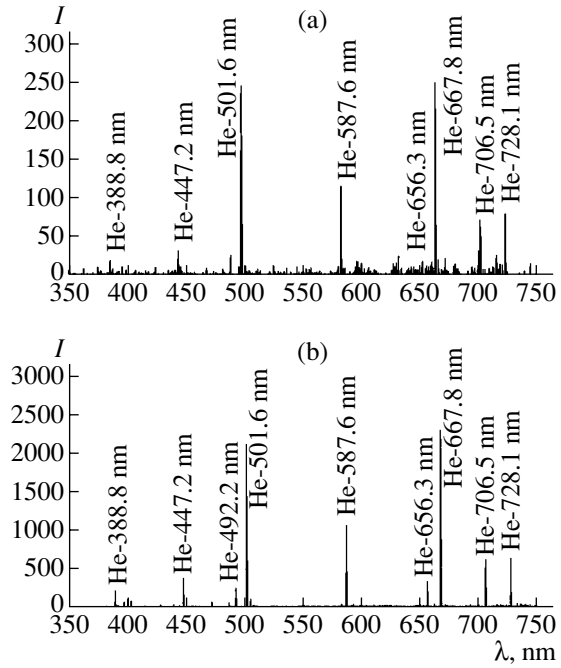


Fig. 8. He spectrum at J_d = 11 mA and P = 130 Pa: (a) in zero magnetic field and (b) in a 0.025-T magnetic field. I is the relative intensity.

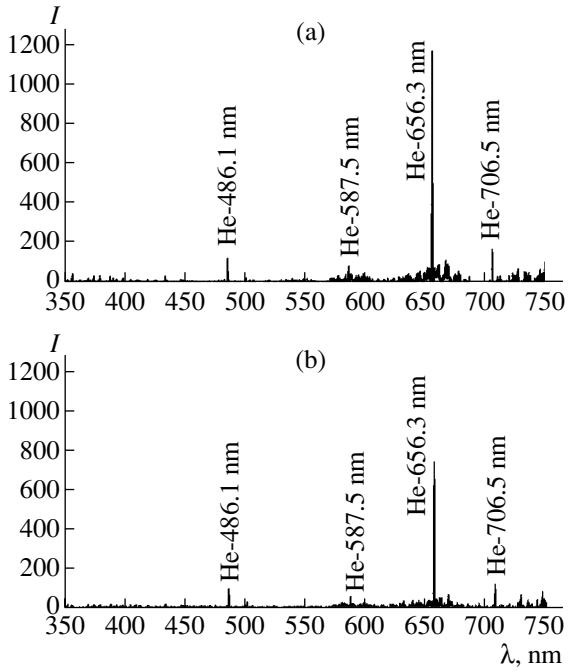


Fig. 9. He spectrum at J_d = 11 mA and P = 180 Pa. Designations (a) and (b) are the same as in Fig. 8.

decrease in the wall current under the action of the magnetic field. The discharge spectrum shown in Fig. 9 also varies only weakly with the magnetic field. The voltage measured between the anode and the cathode and, hence, the electric field strength in this region of the positive column decrease by 5% under the action of the magnetic field at a pressure of 200 Pa. As was shown in [8], at a pressure above 150 Pa in the discharge chamber, a magnetic field as high as 0.08 T did not affect the diffusion of charged particles.

The main difference between our experiments and those described in [3–7, 9] is that we studied the regions of negative glow and the Faraday dark space in the discharge.

Based on the experimental results obtained, we believe that the principal cause of the effect of a longitudinal magnetic field on the discharge is a decrease in the diffusion of charged particles toward the wall under the action of the magnetic field. This decrease leads to a change in the electrical balance of charge carriers and, hence, to a change in the electrical and optical characteristics of the discharge.

Based on the results of [8], we assume that ion and electron diffusion across the magnetic field is expressed by the ambipolar diffusion coefficient D_a,

$$D_a = \frac{D_a(0)}{1 + \omega_e\omega_p\tau_e\tau_p},$$

where $\omega = eB/mc$ is the Larmor frequency, τ is the mean free time of charge carriers, and the subscripts e and p designate electrons and positive ions, respectively.

The calculations performed for our experiments by the formula for the ratio $D_a(B)/D_a(0)$ show that the decrease in the ion current to the wall when the magnetic field increases to 0.05 T at a pressure of 130 Pa in the discharge chamber corresponds to the calculated decrease in the ambipolar diffusion coefficient $D_a(B)$. This result agrees well with the data [8]. When calculating $D_a(B)$, we borrowed the values of ω_e, ω_p, τ_e, and τ_p from [11].

Thus, the general picture of the effect of a longitudinal magnetic field on negative glow and the Faraday dark space of the glow discharge is thought to be the following. The discharge begins to contract near the cathode. Here, the magnetic field is one-tenth of the field between the coils. Electrons begin to move along magnetic lines of force. As the magnetic field increases, the magnetic lines of force become denser and the discharge contracts further. The calculation of this electron motion and the explanation of the discharge contraction are given in [9].

The decrease in the wall loss of charged particles due to the decrease in the ambipolar diffusion coefficient leads to an increase in the electron concentration and temperature in the compressed discharge filament. The velocity distribution function of electrons becomes distorted. This results in a sharp increase in the total luminosity of the compressed discharge and the intensities of some spectral lines (Figs. 7, 8).

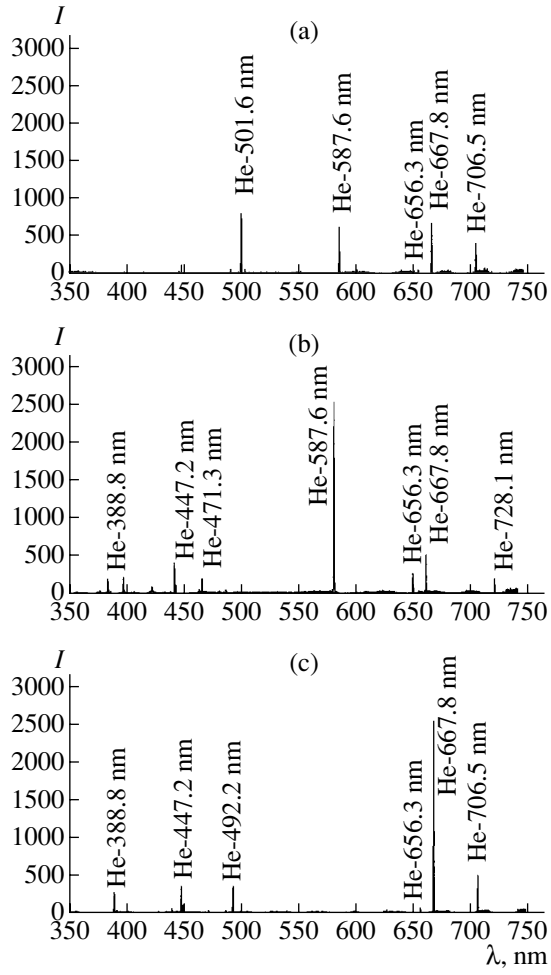


Fig. 10. He spectrum at J_d = 11 mA and P = 100 Pa (50-Hz alternating current): (a) in zero magnetic field and in a magnetic field of (b) 0.025 and (c) 0.05 T.

The decrease in the wall loss of charged particles leads to an increase in the length of negative glow and the Faraday dark space. In our case, the positive column disappears at a pressure lower than 130 Pa in the discharge tube when a magnetic field as low as ~0.005 T is applied. In this case, the whole length of the discharge tube is occupied by cathode glow, dark cathode space, negative glow, and the Faraday dark space.

This picture of the effect of a longitudinal magnetic field on negative glow and the Faraday dark space of the glow discharge is supported experimentally, when we placed a screen of an iron sheet near the magnetic coils from the side of the cathode to shield the magnetic field toward the cathode. In this case, the discharge began to be compressed only behind the screen, and the compression was sharp. The filament was compressed immediately to the diameter required for this magnetic field. However, in the absence of the screen, the filament was compressed gradually from the cathode with increasing magnetic induction. The discharge spectra in the presence of the shield also differed radically from the discharge spectra in the absence of the shield. The intensities of the spectral lines increased by a factor of 3–4, whereas, in the absence of the screen, the intensities of the spectral lines increased by a factor of 10–12. The distortion of the velocity distribution function of electrons in the region of negative glow and the Faraday dark space is significant and supported experimentally at a 50-Hz alternating current. The discharge spectrum in zero magnetic field contains a small number of helium spectral lines (Fig. 10). At a magnetic field of 0.025 T, some lines disappear and some lines appear; at a magnetic field of 0.05 T, some lines again disappear and some lines appear. This behavior is likely to be due only to a strong change in the energy distribution function of electrons.

If we agree with the authors of [12] that atomic excitation for each spectral line is caused only by impacts of the first kind and that stepwise excitation and impacts of the second kind can be neglected, we have to assume that the magnetic field increases the concentration of "beam" electrons, whose energy can excite an atom by impacts of the first kind. We think that the cause of the strong increase in the intensities of the helium spectral lines in a longitudinal magnetic field is the fact that, as the temperature of "base" electrons increases (Fig. 6), the stepwise excitation of the existing long-living metastable levels of helium atoms increases sharply. This results in an increase in the number of excited states that can emit spectral lines.

Note that the authors of [12] also assumed that the effect of a magnetic field on the discharge emission is equivalent to an increase in the number of stepwise excitation acts.

CONCLUSIONS

We have experimentally studied the effect of a longitudinal magnetic field on the regions of negative glow and the Faraday dark space of a glow discharge in helium at a discharge current of 11 mA and a pressure of 100–200 Pa in the discharge chamber.

Using electric probes, we showed that, at a discharge current of 11 mA and a pressure below 130 Pa, a longitudinal electric field is absent (at any rate, it is lower than 0.01 V/cm) in the region of negative glow and Faraday dark space in both the absence and the presence of the magnetic field.

The longitudinal magnetic field is found to change the electrical properties of the discharge: it decreases the total voltage across the discharge and destroys the positive column of the discharge at the given length of the discharge chamber.

When the magnetic field is applied, the wall loss of charged particles decreases. The departure of charged particles to the wall decreases monotonically to a field of ~0.05 T and approximately obeys the law of decreasing of the ambipolar diffusion coefficient (Townsend law).

The effect of the magnetic field on the optical characteristics of the discharge is especially strong. The total emission of the discharge and the intensities of some spectral lines increase by a factor of 10–12.

The detected effects of the longitudinal magnetic field on the glow discharge are explained using the hypothesis that the magnetic field strongly distorts the energy distribution function of electrons, with the fraction of high-energy beam electrons in the total spectrum increasing.

We failed to reveal the effect of the longitudinal magnetic field on the electrical characteristics of the positive column of the discharge at a pressure above 180 Pa in the discharge chamber and a discharge current of 11 mA.

REFERENCES

1. Yu. P. Raĭzer, Teplofiz. Vys. Temp. **29**, 1041 (1991).
2. Yu. P. Raĭzer, Teplofiz. Vys. Temp. **24**, 984 (1986).
3. V. L. Granovskiĭ, *Electric Current in Gases* (Nauka, Moscow, 1971) [in Russian].
4. P. Gill and C. E. Webb, J. Phys. D **10**, 229 (1977).
5. R. J. Bickerton and A. Engel, Proc. Phys. Soc. London, Sect. B **69**, 768 (1956).
6. F. C. Hoh, Phys. Fluids **5**, 22 (1962).
7. F. C. Hoh and B. Lehnert, Phys. Rev. Lett. **7**, 75 (1961).
8. I. A. Vasil'eva and V. L. Granovskiĭ, Radiotekh. Élektron. (Moscow), No. 12 (1959).
9. G. N. Rokhlin, Zh. Éksp. Teor. Fiz. **9** (7) (1939).
10. O. V. Kozlov, *Electrical Probe in a Plasma* (Atomizdat, Moscow, 1969) [in Russian].
11. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1992; Springer-Verlag, Berlin, 1991).
12. V. A. Fabrikant, Doctoral Dissertation (All-Russia Electrotechnical Institute, Moscow, 1939); *Mechanism of Gas Discharge*, in *Tr. Vseross. Élekrotekh. Inst.: Élektron. Ion. Prib.*, Ed. by P. V. Timofeev (Moscow, 1940), Vol. 41 [in Russian].

Translated by K. Shakhlevich

Technical Physics, Vol. 49, No. 9, 2004, pp. 1142–1145. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 39–42.

GAS DISCHARGES, PLASMA

Generation of a Homogeneous Plasma and Wide Ion Beams in a Penning System with a Nonequipotential Cathode

S. P. Nikulin, D. F. Chichigin, and P. V. Tretnikov

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, ul. Komsomol'skaya 34, Yekaterinburg, 620016 Russia

e-mail: nikulin@iep.uran.ru

Received July 17, 2003

Abstract—A study is made of the possibility of controlling the spatial distribution of the degree of ionization in a modified Penning cell with a cathode divided into several sections that are at different potentials. It is shown that an almost uniform plasma can be produced when applying an optimal potential of the control electrode. This enables the generation of wide ion beams in this system. An ion extraction efficiency of several tens of percent is achieved, which is substantially higher than in conventional ion sources. © *2004 MAIK "Nauka/Interperiodica".*

Interest in studies of the characteristics of low-pressure gas discharges and the properties of the plasma generated in them stems, in many respects, from the possibility of employing these discharges for developing sources of charged particle beams. In particular, the production of a homogeneous plasma in large volumes is of great interest. Such a plasma can be used to generate wide beams with a uniform current distribution over the beam cross section.

In [1], it was shown that, in order to produce a homogeneous plasma in a low-pressure discharge, it is necessary that the ionization be nonuniform: the degree of ionization should be lower at the center of the system and higher at the periphery. Therefore, the generation of a homogeneous plasma cannot be achieved if the ionization in the discharge is produced by the plasma electrons. However, in high-current glow discharges with oscillating electrons [2], the greatest contribution to the ionization is made by the fast particles that are generated at the cathode as a result of gamma processes and are accelerated by the cathode voltage drop, which almost coincides with the discharge voltage. Under these conditions, the spatial distribution of the degree of ionization can be controlled by using a nonequipotential cathode, i.e., by dividing the cathode into several sections that are at different potentials. In such a system, the electrons emitted from different cathode sections gain different energies and, consequently, make different contributions to ionization. In particular, to produce a homogeneous plasma, the voltage applied to the central cathode section should be lower than the voltage applied to the periphery of the cathode. In addition, to implement the method proposed, special measures should be undertaken to prevent the rapid mixing of the electrons from different energy groups. This can be achieved by applying an external magnetic field, which reduces the electron mobility in the radial direction. This paper presents the results of experiments on verifying the efficiency of the method proposed.

EXPERIMENTAL SETUP AND THE RESULTS OBTAINED

In experiments, we used a well-known Penning system with a cold cathode [3]. This system has been widely used to produce ion beams. However, the ions were usually extracted through a small emission hole at the axis and, consequently, the beam diameter was small. This was because the discharge plasma was strongly inhomogeneous in the radial direction. In our experiments, we used a modified Penning system (Fig. 1) with hollow cylindrical anode *1* and end cathode *2*. The second end electrode consisted of peripheral

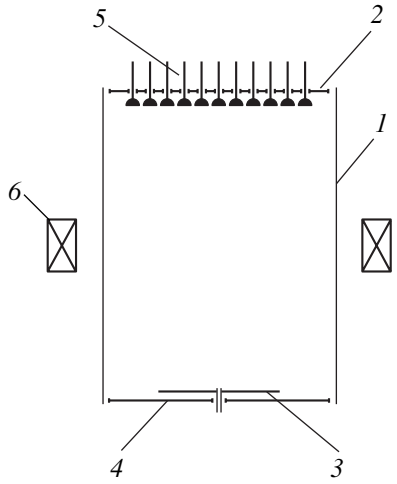


Fig. 1. Modified Penning system with a nonequipotential cathode.

1063-7842/04/4909-1142$26.00

ring *3* and central disk *4*, which was used as a control electrode. The diameter of the control electrode was varied from 30 to 70 mm. The diameter of the gas-discharge system was 135 mm, and its length was varied from 70 to 210 mm. The gas-discharge chamber was embedded in a magnetic field produced by solenoid *6*. The working gas was injected into the system through an axial hole in the central disk and was pumped out through a set of holes in the opposite end cathode. The radial distribution of the ion emission current was measured with the help of a set of probes *5*, arranged at this electrode. The gas pressure was measured by a gauge placed in the vacuum chamber in which the gas-discharge system was installed. In our experiments, the pressure was varied in the range $\sim 10^{-4}$–10^{-3} torr by varying the gas flow rate.

We fixed the discharge current I_d, gas flow rate Q, and magnetic field B and measured the radial distributions of the probe current $I_p(N)$ (where N is the probe number) at different voltages applied to the central disk U_c from a control power source. The voltage at the other cathode sections, which were connected to the negative pole of the discharge power source, was automatically established at a level needed to maintain a given discharge current. A study of the influence of the potential of the central disk on the ion-current distribution showed that this influence was ambiguous. As the voltage applied to the central disk was decreased, the current density profile became smoother and, at some optimal voltage, the distribution became almost uniform. However, a dip in the current density profile then appeared at the center. Figures 2 and 3 show distributions obtained at U_c values close to the optimal one for a disk 30 mm in diameter at two different values of B. A similar behavior was observed for a disk 50 mm in diameter. However, when an even larger, 70-mm-diameter, disk was used, the initial smoothing of the current density profile with decreasing U_c was accompanied by the formation of a sharp maximum at the center of the system. This ambiguous change in the current density profile may be explained as follows: As the voltage U_c decreased, the voltage at all other cathode sections also increased; as a result, the conditions for a self-sustained discharge were satisfied. In this case, however, the energy of the fast particles generated at the discharge periphery increased, which resulted in the enhancement of their transport across the magnetic field toward the anode. When the discharge voltage reached a certain critical value, the ionization rate at the periphery began to decrease because of the increased loss of fast electrons at the anode. At the same time, the ionization rate at the center of the system increased owing to the fast electrons emitted from the central disk not only due to ion-induced electron emission, but also due to the secondary electron emission caused by the electrons originated at the opposite end cathode. As a result, the ionization rate again became maximal near the center of the system. Hence, a uniform plasma density profile

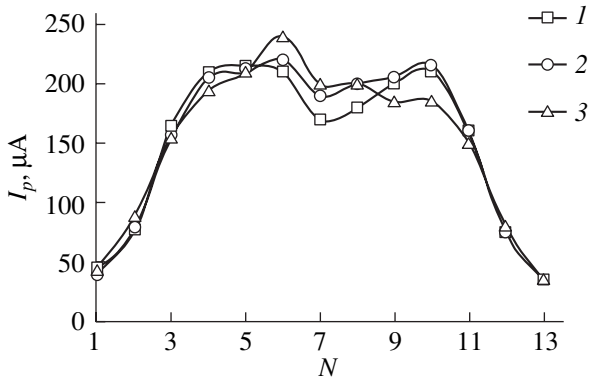


Fig. 2. Radial distribution of the ion current at $B = 11.5$ mT; $Q = 0.3$ cm^3/s; $I_d = 50$ mA; and $U_c =$ (*1*) 95, (*2*) 100, and (*3*) 120 V.

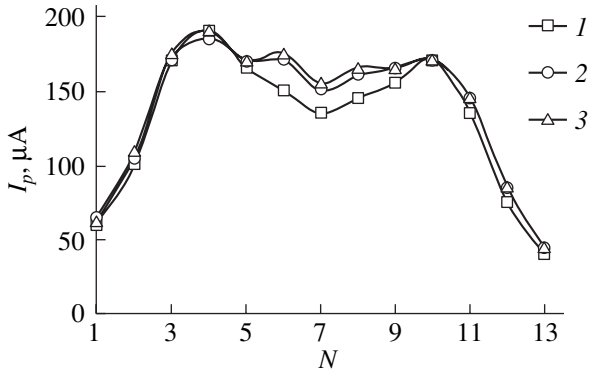


Fig. 3. Radial distribution of the ion current at $B = 14.3$ mT; $Q = 0.3$ cm^3/s; $I_d = 50$ mA; and $U_c =$ (*1*) 80, (*2*) 105, and (*3*) 115 V.

could be achieved by decreasing the potential of the control electrode only if the latter was relatively small in size.

A comparison of the ion current distributions presented in Figs. 2 and 3 shows that the ion current density and, accordingly, the plasma density are lower in a stronger magnetic field. This contradicts the commonly accepted opinion that the particle lifetime in a plasma and, accordingly, the plasma density increase with increasing magnetic field. In connection with this, we studied how the magnetic field influences the discharge conditions. Figure 4 shows the discharge voltage and the current at the central probe as functions of the magnetic field. It turned out that, at a certain critical value of the magnetic field B_1, both the ion emission current and the discharge voltage dropped abruptly. This means that, at this critical field, the discharge conditions also changed abruptly. Oscillograms of the probe current recorded with the help of a C8-17 digital oscilloscope showed that, when the magnetic field was higher than the critical one, the amplitude of the low-frequency ($\sim 10^4$ Hz) modulation of the probe signal reached nearly 100% (in contrast to the case of a subcritical

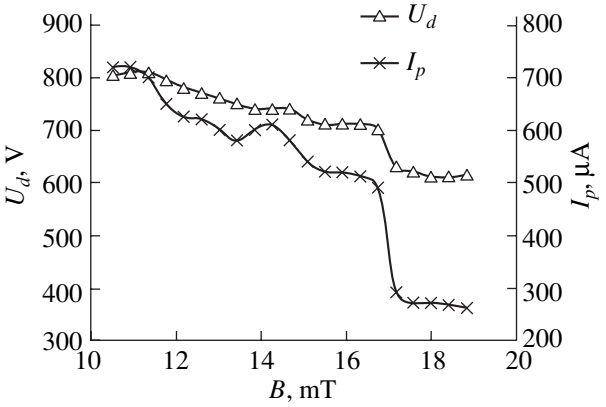


Fig. 4. Discharge voltage U_d and the current I_p at the central probe as functions of B at I_d = 100 mA and U_c = 150 eV.

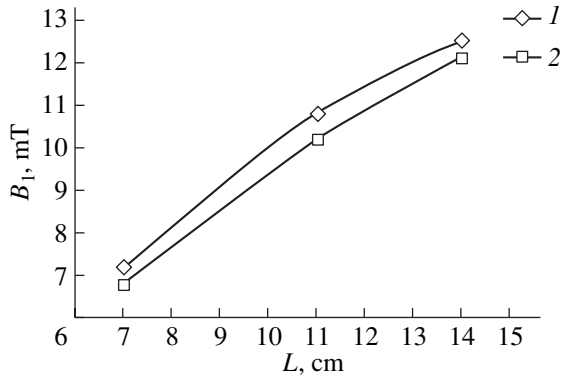


Fig. 5. Critical magnetic field B_1 as a function of the discharge chamber length L at I_d = 50 mA; U_c = 150 eV; and Q = (*1*) 0.2 and (*2*) 0.3 cm^3/s.

magnetic field where this modulation did not exceed several percent). Hence, in a strong magnetic field, we observed the onset of a large-scale low-frequency instability in the gas-discharge plasma. The nature of this instability seems to be similar to that discussed in [3].

Supercritical magnetic fields cannot be used to generate wide ion beams in ion sources with ordinary multiaperture ion optics. A study of the influence of various factors on the magnitude of the critical magnetic field showed that the change in B_1 was most pronounced when we changed the length of the gas-discharge system L and the gas flow rate Q (Fig. 5). It can be seen from Fig. 5 that the value of B_1 increases with increasing L. At large lengths, in the range of magnetic fields under study, the instability could be completely absent. This result can easily be explained taking into account that the onset of instability is closely related to the fact that it is difficult to provide the transport of electrons across the magnetic field onto the anode. It is evident that, as the length of the discharge chamber and, accordingly, the anode area decrease, this transport becomes even more difficult to provide; for this reason, the onset of instability is observed even in a relatively weak field. In contrast, when the discharge chamber length (and, accordingly, the anode area) is large, the transport of electrons to the anode is facilitated and the onset of instability can take place in much stronger magnetic fields. Hence, the upper limit of the magnetic-field operating range increases when using long discharge chambers. When increasing the gas flow rate, we also enhance the electron transport across the magnetic field; as a result, the value of B_1 increases. However, the increase in the gas flow rate in ion sources is limited because of the increase in the probability of breakdown of the acceleration gap.

In order to generate ion beams, we replaced end cathode *2* with an accelerating–decelerating system consisting of three electrodes, each having 163 holes 8 mm in diameter. The accelerating voltage U_a was varied in the range 10–30 kV, and the discharge current was varied in the range 30–100 mA. The ion beam current I_b increased in proportion to the discharge current and depended only slightly on the accelerating voltage. Under the given conditions, this current was 10–30 mA. Hence, the extraction efficiency $\alpha = I_b/I_d$ was independent of the discharge current and was nearly equal to 0.3. When we varied the magnetic field and the gas flow rate, the extraction efficiency also varied only slightly and depended on the geometrical factor only. The extraction efficiency can be estimated by the formula

$$\alpha = 0.5S_h/S, \qquad (1)$$

where S_h is the total area of the emission holes; S is the cross-sectional area of the discharge chamber; and the factor 0.5 is introduced because one-half of the ions produced in a discharge arrives at the upper end, whereas the other half arrives at the lower end of the discharge system.

We also carried out experiments on the generation of low-energy (1–2 keV) ions. Ion beams with this energy are widely used to clean and etch surfaces. In these experiments, the ions were accelerated to a certain prescribed energy by applying a corresponding voltage to the end cathode, rather than by using an ion-optical system. This method of generating ion beams immediately in the cathode sheath was previously used in ion sources based on high-voltage glow discharges. These systems, however, did not find wide application. The reason is that the high pressures required to maintain discharges in such systems are unacceptable for many processes of ion-plasma technology. In addition, in discharges of this kind, the main characteristics (such as the current, voltage, and working gas pressure) are interrelated; as a result, it is difficult to control the ion-beam parameters. In our system, which is based on the Penning discharge with a nonequipotential cathode, the working pressure is several orders of magnitude lower. Our experiments showed that the discharge did not lose its stability when an increased voltage corresponding to the required ion energy was applied to the end cathode. The current of ions arriving at the end cathode, which

served as a workpiece, was controlled by the discharge power source inserted in the circuit between the anode and the peripheral ring. Here, we used the circumstance that, over a wide range of the discharge parameters, the currents at both cathodes were nearly equal to one another; therefore, an increase in the discharge current automatically resulted in an increase in the current at the target. It should be noted, however, that, in these experiments, the current in the circuit of the end cathode somewhat increased with increasing accelerating voltage. This is probably due to an increase in the coefficient of ion-induced electron emission. The radial profile was controlled, as before, by applying a lower voltage to the central disk as compared to that applied to the ring. In these experiments, the voltage U_c was decreased by connecting the central disk 30 mm in diameter to the discharge power source through a high-resistance resistor. Here, we used the circumstance that, for a disk of this size, the optimal value of the potential was close to the floating potential. When the accelerating voltage was increased, we did not observe a substantial distortion of the radial profile of the ion current.

CONCLUSIONS

The use of a nonequipotential cathode in a Penning system makes it possible to efficiently control the spatial distribution of the plasma density and to achieve a nearly uniform distribution of the ion-emission current in a device with a small-size control electrode. The maximum attainable ratio of the ion-beam current to the discharge current (~30% with an ion optics and ~50% with the use of acceleration in the cathode sheath) in this system is several times higher than that achieved in conventional sources. Accordingly, the energy efficiency increases substantially. The results of our studies show that Penning systems with nonequipotential cathodes show promise as plasma emitters in ion sources generating wide ion beams and also in systems for ion cleaning and etching.

ACKNOWLEDGMENTS

This work was supported in part by the Russian foundation for Basic Research (project no. 02-02-17860) and the RF Ministry of Industry, Science, and Technologies (grant no. 40.030.11.1126).

REFERENCES

1. S. P. Nikulin and S. V. Kuleshov, Zh. Tekh. Fiz. **70** (4), 18 (2000) [Tech. Phys. **45**, 400 (2000)].
2. A. S. Metel', Zh. Tekh. Fiz. **55**, 1928 (1985) [Sov. Phys. Tech. Phys. **30**, 1133 (1985)].
3. M. D. Gabovich, *Physics and Technology of Ion Plasma Sources* (Atomizdat, Moscow, 1972) [in Russian].

Translated by N. Larionova

Technical Physics, Vol. 49, No. 9, 2004, pp. 1146–1149. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 43–46.

GAS DISCHARGES, PLASMA

Arc Discharge Sputtering of Schungites

L. A. Aleshina, V. I. Podgornyi, G. B. Stefanovich, and A. D. Fofanov

Petrozavodsk State University, Petrozavodsk, 185014 Russia

e-mail: podgvlad@mainpgu.karelia.ru

Received November 3, 2003

Abstract—The schungite structure is shown to be modified in an arc discharge in an inert gas atmosphere. As a result of the direct effect of the arc discharge, some portion of the schungite material transforms locally into rhombohedral graphite. The spectral dependence of the optical transmission of the solutions of graphite soot and schungite soot in toluene is measured. X-ray diffraction patterns of the schungite soot and graphite soot contain a broad maximum at small scattering angles, which indicates the presence of fullerene molecules.

INTRODUCTION

Schungite is a natural mineral containing from 1 to 98% carbon in a noncrystalline form, microelements, a mineral component, and a small amount of a bituminous organic component. Carbon exists in schungites in the form of micro- and nanoclusters—globules ~10 nm in size [1, 2]. The similarity of the subspherical graphite-like layers of the schungite carbon and fullerene structures suggests that fullerenes can rather easily be produced from this mineral. Moreover, as follows from some works [3], it contains such structures in its initial form. Therefore, it is interesting to verify the possibility of producing fullerenes from schungites by sputtering them in an arc discharge according to the well-known technique [4]. The purpose of this work is to study schungites of two types; samples of one type were produced from rock containing 70% bound carbon (samples Sh1), and the other samples were produced from rock containing 98% C (samples Sh2).

EXPERIMENTAL

Cleaved schungite samples 2–6 mm in thickness were fixed on a graphite substrate, which was connected to the positive potential of a 0–40 V dc voltage source. The cathode was a sharp-nosed ChDA graphite electrode made 6 mm in diameter. As a buffer gas, we used high-purity argon and helium. An arc was initiated by resistance heating of the cathode tip during brief contact with a sample. The interelectrode gap in the operating mode was varied from 0.5 to 2 mm.

The voltage of arcing in an argon atmosphere was about 20 V, and that in a helium atmosphere was about 30 V. The arc current was varied in the range 40–100 A, and the gas pressure, from 1×10^4 to 4×10^4 Pa.

The electrodes were arranged horizontally at the end faces of a water-cooled cylindrical chamber. One of the electrodes could be moved along its axis. The chamber was placed inside a vacuum bell jar having two peepholes. This design allowed us to observe arc initiation, the motion of a cathode spot, the interelectrode gap, and the arc shape. The chamber was preliminarily evacuated to a residual pressure of $p \cong 1$ Pa with a backing pump.

The sputtering time was varied in the range from 5 to 60 min. Schungite soot was mechanically removed from the chamber walls and flooded with toluene. After storing the soot in toluene for several days, we measured the spectral dependence of the optical transmission of the mixture with an SF-26 spectrophotometer. As a reference solution, we used a ruby toluene–graphite soot infusion; the graphite soot was produced under the same arc conditions after a sample was replaced by graphite.

X-ray diffraction analysis of the samples was performed on a DRON-4 apparatus with FeK_α and CuK_α radiation. As a monochromator, we applied a crystal of pyrolytic graphite.

RESULTS AND DISCUSSION

Heating of the samples is found to result in their cleavage into several low-strength plates about 1 mm in thickness, whose planes are predominantly parallel to each other. When the temperature increases rapidly, some plates fly away from a sample, which strongly changes its dimensions. An arc discharge produces craters with a depth of 5 mm in the samples; the craters can go through some plates and, to some extent, can fasten them together. The graphite cathode is not sputtered; it even gains weight due to sputtering of the anode mate-

1063-7842/04/4909-1146$26.00

rial. It is seen that, after the electrodes (especially the cathode tip) heated to a high temperature are moved apart, the tip acquires a buildup having the same diameter and a length of 1–2 mm. For the electrodes to be separated, a certain mechanical force must be applied to them. Obviously, during resistance heating to high temperatures, some portion of the sample material is welded to the cathode tip. Thus, in essence, the arc initiated under these conditions glows between two schungite electrodes.

Upon the following arcing, especially at a sufficiently small interelectrode gap ($d \le 0.5$ mm, i.e., in the so-called contact-arc mode) and a low discharge current ($I \le 50$ A), the buildup increases in size. Long-term arcing in this mode results in a buildup in the form of an irregular hemisphere that has needle branches growing in all directions and is completely located in the crater of a sample. Figure 1 shows the X-ray diffraction pattern of one of such buildups that is taken from sample Sh2 and the X-ray diffraction pattern of schungite. X-ray line profile analysis was performed by the technique [5] using the atomic coordinates of hexagonal and rhombohedral graphite modifications [6]. The analysis shows that the lattice parameters of the sample under study correspond to rhombohedral graphite. In other words, the buildup of rhombohedral graphite is formed on the graphite electrode during arc sputtering of schungite. It should be noted that the cathode material was hexagonal graphite.

X-ray diffraction studies of the schungite samples after the arc discharge indicate that, within the limits of experimental error, the X-ray diffraction patterns of the bottom and the lateral surfaces of craters coincide with the X-ray diffraction pattern of the cylindrical buildup given in Fig. 1 and correspond to rhombohedral graphite. We assume that the arrangement of graphite networks in the initial schungite favors the formation of this allotropic carbon modification.

Therefore, the high-temperature arc treatment of schungite results in its graphitization. It is known that nanostructures should be formed on graphite in the cathode deposit under these conditions [7, 8]. However, the X-ray diffraction patterns of the buildup do not reveal scattering by nanostructures (see Fig. 1).

In the arcing mode with a large interelectrode distance ($d \ge 1$ mm), the buildup does not gain weight when the discharge current increases to 80–100 A. It is visually found that, under these conditions, the cathode can be completely cleaned of the sputtered anode material. This finding supports, to some extent, the absence of nanotubes in our case; these tubes are known to be very stable structures [7].

Figure 2 shows the results of measuring the optical transmission of the soot infusions of the samples as compared to the transmission of pure toluene. Curve *1* corresponds to samples Sh1; curve *2*, to samples Sh2; and curve *3*, to ruby toluene–graphite soot infusion. As is seen from curve *3*, the colored infusion has two

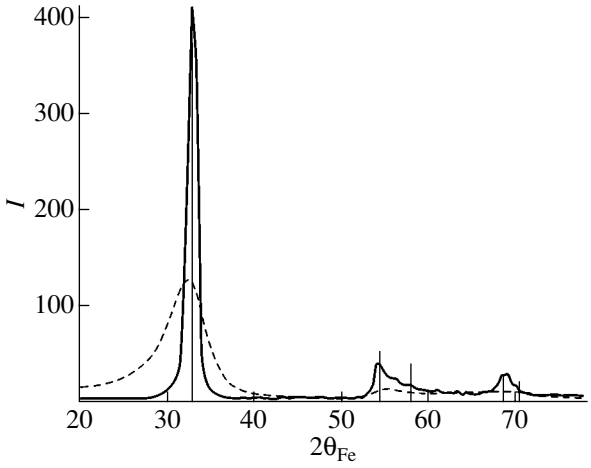


Fig. 1. X-ray diffraction patterns of (dashed line) sample Sh2 in the initial state and (solid line) the buildup on the graphite cathode and (vertical bars) bar diagram of rhombohedral graphite. *I*, counts.

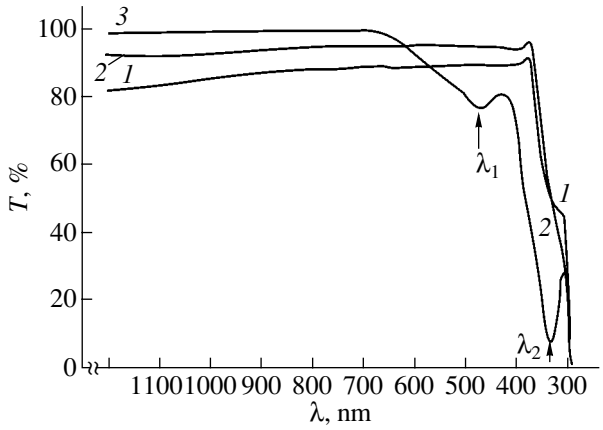


Fig. 2. Dependence of the optical transmission of the soot infusion on the wavelength: (*1*) samples Sh1, (*2*) samples Sh2, and (*3*) ruby toluene–graphite soot infusion.

absorption maxima at $\lambda_1 = 470$ nm ($E_1 = 2.6$ eV) and $\lambda_2 = 335$ nm ($E_2 = 3.7$ eV).

The strong absorption of the fullerene-containing infusion of toluene at a wavelength coinciding with λ_2 was earlier detected in [9]. Therefore, we can assert that molecules of C60 and C70 fullerenes dissolved in toluene are present in our samples of graphite soot.

As follows from the energies of the absorption maxima at λ_1 and λ_2, these molecules are unlikely to be related to, for example, the decomposition of clusters or dimers of C60 and C70 molecules or the excitation of any vibrations in these structures. We think that these molecules can result from electronic processes.

Curve *1* has no specific features near the maxima, whereas curve *2* has a small "step" at a wavelength that virtually coincides with λ_2. This coincidence may indi-

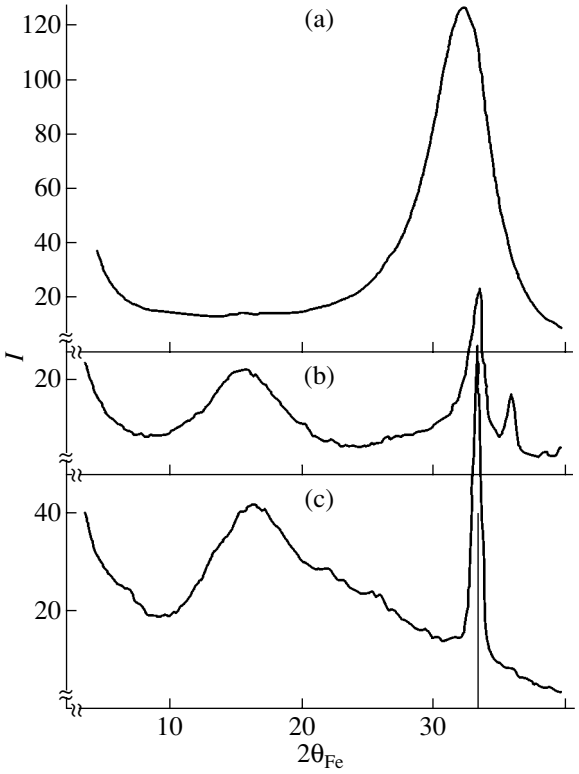


Fig. 3. X-ray diffraction patterns of (a) the initial schungite, (b) the soot deposited during the evaporation of schungite in an arc, and (c) the soot deposited during the evaporation of graphite in an arc. *I*, counts.

cate the presence of structures that are close to fullerenes in the schungite Sh2 soot. Therefore, we also studied the X-ray diffraction patterns of the soot deposited onto the chamber walls during arc sputtering of the schungite samples.

The results of these studies are given in Fig. 3. Curve *1* corresponds to the initial schungite Sh1 powder; curve *2*, to the schungite soot after arc sputtering; and curve *3*, to the graphite soot deposited under the same conditions. The initial schungite Sh1 sample is seen to have a broad maximum in an angular range that virtually coincides with the positions of similar maxima for charcoal and anthracite powders [6]; this maximum differs from them only in the half-width. In the schungite soot (curve *2*), an additional pronounced symmetric maximum appears in the angular range $2\theta = 10°–30°$. In the case of the graphite soot (curve *3*), this maximum becomes sharply asymmetric and its intensity almost doubles.

Since the graphite soot (unlike the schungite soot) contains molecules of C60 and C70 fullerenes, we can assume that the asymmetric part of the maximum on curve *3* is caused by the presence of these chaotically oriented molecules. From this standpoint, the symmetric maximum on curve *2* and the corresponding part of the maximum on curve *3* can be related to the presence of fullerene molecules with a large number of atoms, e.g., C84, in the schungite soot and graphite soot. In other words, an increase in the number of atoms in a fullerene molecule leads to a shift in the maximum toward smaller scattering angles.

Thus, the soot deposited during the arc-induced evaporation of graphite contains molecules of heavier fullerenes apart from C60 and C70 fullerenes. The soot deposited during the evaporation of schungite in the arc is likely to have only higher fullerenes.

The results obtained suggest two basic methods of producing fullerenes in schungite soot. First, they can be formed as a result of coiling of globular carbon followed by the evaporation of the formed molecules from schungite. Second, they can be assembled in the arc plasma by one of the mechanisms that are discussed, in particular, in [8], with the necessary preliminary stage of this processes being the transition of the schungite carbon into the rhombohedral phase. This stage can also occur in the case of sputtering of graphite in an arc.

CONCLUSIONS

(1) The structure of schungite has been shown to be modified in an arc discharge in an inert gas atmosphere.

(2) As a result of the direct effect of the arc discharge, some portion of the schungite material transforms locally into rhombohedral graphite.

(3) X-ray diffraction patterns of the schungite soot and graphite soot contain a broad maximum at small scattering angles, which indicates the presence of fullerene molecules.

ACKNOWLEDGMENTS

We thank Professor V.V. Vapirov for his consultation and assistance in the spectrophotometric studies, Professor V.I. Sysun for useful comments regarding the physics of arc discharge, and Professor Yu.V. Voĭtekhovskiĭ for providing the atomic coordinates in fullerene molecules.

This work was supported by the Ministry of Education of the Russian Federation (project no. PD02-1.2-183) and the Civilian Research and Development Foundation (CRDF) (grant no. PZ-013-02).

REFERENCES

1. *Karelian Schungites and the Ways of Their Complex Harnessing*, Ed. by V. A. Sokolov and Yu. K. Kalinin (Kareliya, Petrozavodsk, 1975) [in Russian].

2. *Organic Substance of Karelian Schungite-Bearing Minerals*, Ed. by M. M. Filippov (KNTs RAN, Petrozavodsk, 1994) [in Russian].
3. A. Z. Zaĭdenberg *et al.*, Zh. Fiz. Khim. **70**, 107 (1996).
4. A. V. Eletskiĭ and B. M. Smirnov, Usp. Fiz. Nauk **165**, 977 (1995) [Phys. Usp. **38**, 935 (1995)].
5. L. A. Aleshina and S. V. Loginova, Kristallografiya **47**, 460 (2002) [Crystallogr. Rep. **47**, 415 (2002)].
6. T. G. Shumilova, *Diamond, Graphite, Fullerene and Other Carbon Modifications* (Ural. Otd. RAN, Yekaterinburg, 2002) [in Russian].
7. A. V. Eletskiĭ, Usp. Fiz. Nauk **167**, 945 (1997) [Phys. Usp. **40**, 899 (1997)].
8. Yu. E. Lozovik and A. M. Popov, Usp. Fiz. Nauk **167**, 751 (1997) [Phys. Usp. **40**, 717 (1997)].
9. N. S. Anikina *et al.*, in *Proceedings of the Conference "Fullerenes and Fullerene-Like Structures," Minsk, 2000.*

Translated by K. Shakhlevich

Technical Physics, Vol. 49, No. 9, 2004, pp. 1150–1157. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 47–54.

GAS DISCHARGES, PLASMA

Mechanism for the Streamer Propagation toward the Anode and Cathode due to Background Electron Multiplication

S. I. Yakovlenko

Prokhorov Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 119991 Russia

Received January 8, 2004

Abstract—A simple mechanism for the propagation of an ionization wave in a dense gas due to the multiplication of background electrons in a nonuniform electric field is proposed. The mechanism does not depend on the sign of the field projection onto the streamer propagation direction. The streamer propagation is caused by the enhancement of the electric field at the streamer head. It is shown that, in a prebreakdown field, the intense multiplication of electrons takes place in both electropositive and electronegative gases. The prebreakdown multiplication can provide a fairly high density of background electrons; this allows one to treat the background as a continuous medium when considering streamer propagation as a multiplication wave. The initial ionization is enabled by the natural background of ionizing radiation and cosmic rays. An analytical expression for the velocity of the ionization front is obtained based on a simple equation for the multiplication of background electrons. This expression is in good agreement with numerical simulations performed within both a simple model of background electron multiplication and a more comprehensive drift–diffusion model. In particular, the drift–diffusion model predicts the propagation of the ionization front from a small-radius anode to the cathode due to the multiplication of background electrons. The velocity of the ionization wave front is calculated as a function of the electric field at the streamer head for helium, xenon, nitrogen, and sulfur hexafluoride. It is shown that some features of streamer propagation (e.g., its jerky motion) can be related to the recently found nonmonotonic dependence of ionization frequency on the electric field. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is well known [1–7] that a streamer forms from an electron avalanche in a dense gas when the electron density in an avalanche increases to such high values that the newly created plasma begins to enhance the external electric field. The mechanism for streamer propagation is still poorly understood, although this issue has long been studied and there are a lot of papers devoted to numerical simulations of streamers (see, e.g., [8–14] and the literature cited therein).

The propagation of a streamer toward the anode is usually attributed to the drift motion of electrons at the streamer head. However, the existence of cathode-directed streamers causes surprise. At present, a photon mechanism proposed by Meek, Loeb, and Raether [1–7] is usually invoked to describe a cathode-directed streamer. According to this mechanism, the streamer head emits photons that ionize the gas in front of it. The photoionization events give rise to avalanches directed toward the head. Overlapping with one another, these avalanches enable the streamer propagation. This theory was further developed by Firsov and Lozanskiĭ [6]. In particular, they assumed that the transfer of the line radiation and the subsequent associative ionization of the excited states play an important role in the streamer propagation.

A number of fundamental drawbacks of the photon theory compel one to search for other mechanisms for streamer propagation. For example, in [15, 16], a nonphoton mechanism related to Langmuir plasma oscillations was proposed. However, these ideas did not take the form of a completed theory. This is why, though being constantly criticized, the photon hypothesis still remains the most commonly accepted model. Probably, this is the only reason that there has not yet been another clearly formulated model for the propagation of the ionization region toward the cathode.

In this paper, a rather simple nonphoton mechanism is proposed that accounts for the streamer propagation toward both the anode and the cathode. Within this model, the streamer propagation is explained by the exponential multiplication of the background electrons (that are always present in a gas) in a nonuniform electric field, rather than by the electron drift or radiation transfer.

DENSITY OF BACKGROUND ELECTRONS IMMEDIATELY BEFORE THE STREAMER BREAKDOWN

1. Electropositive Gases

It is well known that the natural background of ionizing radiation and cosmic rays amounts to ~10 μR/h, which corresponds to an electron production rate of $G_i \sim 6\ \text{cm}^{-3}\ \text{s}^{-1}$. In atmospheric-pressure noble gases (i.e., at a neutral particle density of $N \approx 2.4 \times 10^{19}\ \text{cm}^{-3}$),

1063-7842/04/4909-1150$26.00 © 2004 MAIK "Nauka/Interperiodica"

atomic ions convert into molecular ones over a time of ~10 ns, whereas the dissociative recombination rate constant is $k_d \sim 10^{-7}$ cm^3/s. Consequently, the background electron density N_{e0} under equilibrium conditions, when $G_i = k_d N_{e0}^2$, is

$$N_{e0} = (G_i/k_d)^{1/2} \sim 10^3 \text{ cm}^{-3}. \quad (1)$$

The multiplication of electrons in the external electric field leads to an increase in their background density even without breakdown. Let us illustrate this using a simple model for the electron balance between plane electrodes:

$$\frac{dN_e}{dt} = \left(\nu_i - \frac{u_{d,e}}{d}\right)N_e - k_d N_e^2, \quad N_e(0) = N_{e0}. \quad (2)$$

Here, ν_i is the ionization frequency, $u_{d,e}$ is the electron drift velocity, d is the distance between the electrodes, and N_{e0} is determined by formula (1). It is assumed that a constant field is instantaneously switched on at the time $t = 0$. The solution to Eq. (2) is

$$N_e(t) = \frac{N_{e0}\exp(\nu t)}{1 + (N_{e0}k_d/\nu)(\exp(\nu t) - 1)},$$

where the effective frequency $\nu \equiv \nu_i - u_{d,e}/d$ can be either positive or negative.

The ionization frequency can be written as the product $\nu_i = \alpha_i u_{d,e}$ of the Townsend coefficient $\alpha_i(E, p)$ and the electron drift velocity $u_{d,e}(E/p)$ [7]. The Townsend coefficient can be expressed as $\alpha_i(E, p) = p\xi(E/p)$, where $\xi(E, p)$ is a function characteristic of a given gas, E is the electric field strength, and p is the gas pressure. The calculated values of these quantities (taken from [17–19]) are shown in Fig. 1.

Multiplication takes place when $\nu_i > u_{d,e}/d$, i.e., at $\alpha_i d > 1$. For plane electrode geometry, when the voltage between the electrodes is $U = Ed$, it follows from the condition $\alpha_i d = 1$ that $pd\xi(U_{cr}/pd) = 1$. The latter equation determines the "escape curve" that separates the parameter region corresponding to efficient electron multiplication from the parameter region where electrons escape from a discharge without multiplication (see [17–19] for details). Each gas has its own escape curve. The calculated escape curves for different gases are shown in Fig. 2.

The condition $U > U_{cr}(pd)$ is still insufficient for sustaining a steady-state discharge. According to the commonly accepted views, the ion density should be high enough to enable a sufficiently intense electron production due to secondary electron emission (see, e.g., [7]). Hence, the occurrence of breakdown corresponds to the Paschen curve whose right branch lies somewhat above the escape curve (see [17–19] for details). Nevertheless, at atmospheric pressure, an intense multiplication of electrons begins under the condition $U > U_{cr}(pd)$. In the model under consider-

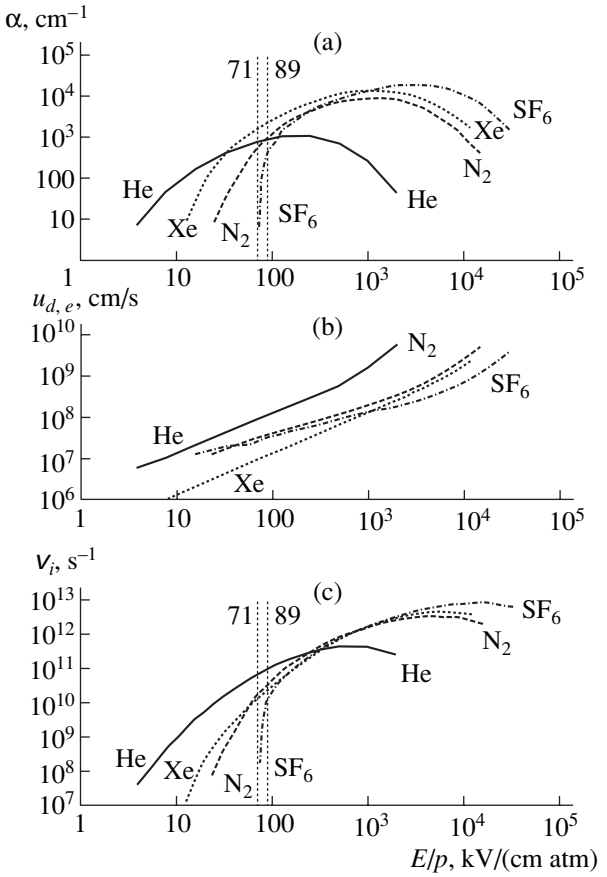


Fig. 1. (a) Townsend multiplication coefficients, (b) electron drift velocities, and (c) ionization frequencies vs. reduced electric field E/p for different gases at atmospheric pressure (by the data from [17–19]).

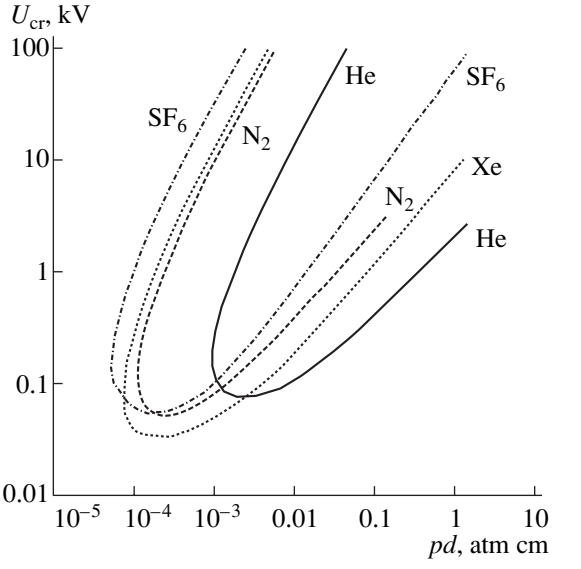


Fig. 2. Escape curves $U_{cr}(pd)$ separating the parameter region corresponding to electron multiplication from the parameter region where electrons escape from a discharge without multiplication for different gases (by the data from [17–19]).

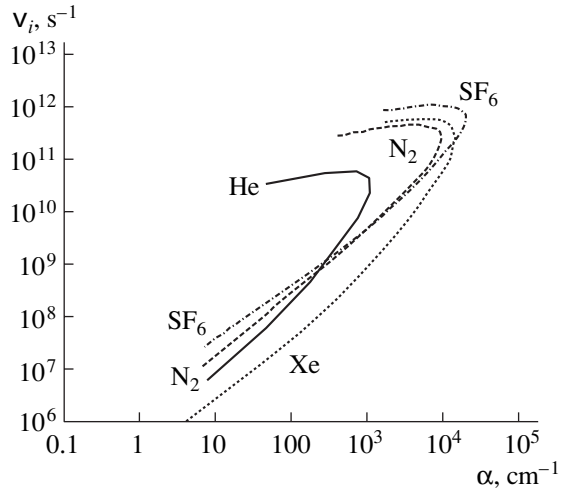


Fig. 3. Ionization frequency vs. Townsend multiplication coefficient for different gases at atmospheric pressure (by the data from [17–19]).

ation, the electron density is limited to the value (for $d \approx 1$ cm)

$$N_{e\max} = \nu_i/k_d \sim 10^{14}\ \text{cm}^{-3}. \quad (3)$$

Actually, under the above conditions, the value $N_{e\max}$ cannot be reached because the external field is shielded by the plasma. It is because of the high ionization frequency $\nu_i \propto N \propto p$ and, accordingly, the high maximum electron density $N_{e\max}$ that a streamer is formed at high pressures. Shielding begins to play an important role at $4\pi e d N_{e\max} > E$.

It can be seen from Fig. 3 that, at atmospheric pressure and $d \approx 1$ cm, the condition $\alpha_i d = 1$ (or $U = U_{cr}$) is satisfied at electric fields at which the ionization frequency is already fairly high. Consequently, the background electron density at which the distance between the electrons is much less than the size of the streamer head is reached over a relatively short time.

Indeed, taking, for example, $\alpha_i d = 2$ (or $\nu_i = u_d/2d$), we find that the time required to reach a certain background electron density $N_{e\text{bg}} \ll N_{e\max}$ is

$$\tau = 2\nu_i^{-1}\ln(N_{e\text{bg}}/N_{e0}).$$

It can be seen from Fig. 3 that, for $N_{e\text{bg}} \sim 10^6$ cm^{-3} (which corresponds to a mean interelectron distance on the order of 0.1 mm), we have $\nu_i \sim 10^7$ s^{-1} and, accordingly, $\tau \sim 1$ μs. This estimate corresponds to the minimum fields at which the multiplication of the background electrons is still feasible. In higher fields, the background is formed over a much shorter time (before the formation of a streamer). The reason is that, to enable the streamer formation, the electron density must be high enough to ensure the condition $\nu_i d/u_{d,e} \sim 20$ [1–3]. To provide a background electron density that is much less than the electron density in the streamer head, it is sufficient that $\nu_i d/u_{d,e} \approx 5$. For this reason, the background is formed at lower fields.

Thus, an increase in the voltage above the escape curve results in a sharp increase in the background electron density throughout the entire volume occupied by the electric field until a plasma bunch shielding the field forms in a certain spatial region.

2. Electronegative Gases

In electronegative gases in the absence of an electric field, the electron density is much lower than that determined by formula (1). Electrons attach to neutrals due to both pair collisions while degrading over energy through the energy range of a few electronvolts and three-body collisions at room temperature. At the same time, the density of negative and positive ions is ~10^3 cm^{-3}.

When considering electron multiplication in an electronegative gas within model (2), it is necessary to take into account electron attachment and the fact that the electron density is zero at the initial time:

$$\frac{dN_e}{dt} = \left(\nu_i - \frac{u_{d,e}}{d}\right)N_e + G_i, \quad N_e(0) = 0. \quad (4)$$

Here, in contrast to the case of electropositive gases, the quantity $\nu_i \equiv \nu_{\text{mult}} - \nu_{\text{att}}$ accounts for both multiplication and attachment of electrons (ν_{mult} and ν_{att} are the multiplication and attachment frequencies, respectively). The initial stage of multiplication, when $N_e \ll N_{e\max}$, is considered. The solution to Eq. (4) is

$$N_e(t) = \frac{G_i}{\nu}(\exp(\nu t) - 1), \quad \nu \equiv (\nu_i - u_{d,e}/d). \quad (5)$$

In electronegative gases, the dependence of ν_i on the reduced electric field passes through zero at a certain E/p value. For example, in SF$_6$, we have $\nu_i = 0$ at E/p = 71 kV/(cm atm) [19]. It is experimentally shown, however, that breakdown occurs at a somewhat higher field, E/p = 89 kV/(cm atm) [7], at which the ionization frequency is substantially increased (Fig. 1).

Taking $\alpha_i d = 2$ (or $\nu_i = u_d/2d$), we find that the time during which a certain electron density $N_{e\text{bg}} \gg G_i/\nu_i$ is reached is

$$\tau = 2\nu_i^{-1}\ln(G_i N_{e\text{bg}}/\nu).$$

Let us take $N_{e\text{bg}} \sim 10^6$ cm^{-3} and use the data for SF$_6$ and E/p = 77 kV/(cm atm) from Fig. 3. We then have $\nu_i \sim 2 \times 10^9$ s^{-1}; $G_i/\nu_i \sim 10^{-7}$ cm^{-3}; and, accordingly, $\tau \sim$ 30 ns.

Thus, as in the case of electropositive gases, rapid electron multiplication takes place at prebreakdown electric fields. However, electron multiplication in electronegative gases begins at significantly higher fields than in electropositive gases.

PROPAGATION OF THE FRONT OF BACKGROUND ELECTRON MULTIPLICATION IN A NONUNIFORM ELECTRIC FIELD

1. Simple Model

Let a small plasma body with a sufficiently high density (such that the electric field is zero inside it) be formed between plane electrodes. Let the plasma body be surrounded by a gas with a very low degree of ionization. In other words, the electron density outside the plasma body is many orders of magnitude lower than inside it.

In this paper, we consider an ionization propagation mechanism related to the exponential multiplication of the low-density background electrons in a nonuniform electric field, rather than to the electron drift or radiation transfer. The higher the electric field at a given spatial point, the higher the rate of electron multiplication. The field is concentrated at the streamer head. Therefore, electron multiplication is most intense near the streamer head, where the field is the highest. The increase in the electron density is accompanied by the shielding of the electric field and the shift of the plasma boundary. Hereinafter, this phenomenon will be called a background multiplication wave.

Let us describe the multiplication wave using the following simple model. We ignore the electron drift and define the boundary between the streamer and the neutral gas as the surface at which the plasma density reaches a certain critical value N_{cr} that corresponds to the complete shielding of the electric field. The electron density at any spatial point with a radius-vector $\mathbf{r}$ is then given by the expression

$$N_e(\mathbf{r}, t) = \begin{cases} N_0 \exp[\nu_i(E(\mathbf{r}))t] & \text{for } N_0 \exp[\nu_i(E(\mathbf{r}))t] < N_{cr} \\ N_{cr} & \text{for } N_0 \exp[\nu_i(E(\mathbf{r}))t] \geq N_{cr}, \end{cases} \quad (6)$$

where N_0 is the background plasma density. It is clear that, within model (6), the propagation direction of the ionization wave does not depend on the sign of the projection of the field on the propagation direction because the ionization frequency is determined by the absolute value of the electric field.

2. Velocity of the Multiplication Front

The coordinates of the front of the multiplication wave are determined by the points at which the electron density reaches its critical value. Let us consider how the coordinate $z(t)$ of one of the front points varies along the normal to the front. The dependence $z(t)$ is implicitly determined by the expression

$$\nu_i(E_0(z(t)))t = Ln, \quad Ln \equiv \ln(N_{cr}/N_0), \quad (7)$$

where $E_0 = E(z(0))$ is the electric field at the front surface.

Generally, the quantity Ln, together with N_{cr}, is a function of E_0. However, we ignore this dependence because it is logarithmic. Taking the derivative of expression (7), we obtain

$$u_{fr} = \frac{dz}{dt} = \nu_i \left[\left(\frac{d\ln\nu}{d\ln E} \left| \frac{-\nabla E}{E} \right| \right)_{E = E_0} Ln \right]^{-1}. \quad (8)$$

If we approximate a piece of a surface near the streamer head with a sphere of radius r_0, then we obtain $|-\nabla E/E|_{E = E_0} = 2/r_0$. Accordingly, we have

$$u_{fr} = \nu_i r_0 \left[\left(\frac{d\ln\nu}{d\ln E} \right)_{E = E_0} 2Ln \right]^{-1}. \quad (9)$$

As was mentioned above, the ionization frequency can be written as the product of the Townsend coefficient $\alpha_i(E, p) = p\xi(E/p)$ and the electron drift velocity $u_{d,e}(E/p)$: $\nu_i = \alpha_i u_{d,e}$. It is important that the most significant ionization characteristics are functions of the reduced electric field E/p. Therefore, the propagation velocity of the streamer head can be expressed through the functions of E_0/p,

$$u_{fr} = \nu_i r_0 / \zeta(E_0/p),$$
$$\zeta(E_0/p) = 2\mathrm{Ln}\left(\frac{d\ln(u_{de}(E/p)\xi(E/p))}{d\ln(E/p)} \right)_{E/p = E_0/p}, \quad (10)$$

which are unique for a given gas.

In [17–19], it was found that the dependence $\nu_i(E, p)$ has a maximum at a certain field strength $E/p = (E, p)_{cr}$. This should result in some specific features of the propagation of the ionization front that will be discussed below.

3. Front Velocity in Helium and Xenon

Let us consider in more detail the propagation of the ionization front in helium and xenon, because the ionization–drift characteristics of these gases were thoroughly investigated numerically and simple approximations were obtained in [17–19]. For helium [17], we have

$$\xi(x) = 5.4(\mathrm{torr})^{-1} \exp(-(14/x)^{1/2} - 1.5 \times 10^{-3} x),$$
$$u_{de} = 10^6 \text{ cm/s}. \quad (11)$$

Substituting expression (11) into formula (10), we obtain

$$u_{fr} = \nu_i r_0 / \zeta(x),$$
$$\zeta(x) = 2Ln(1 + 1.87x^{-1/2} - 1.5 \times 10^{-3}), \quad (12)$$

where $x = (E_0/p)$ V/(cm Torr). For helium, we have $(E_0/p)_{cr} \approx 720$ V/(cm Torr). For xenon, the following

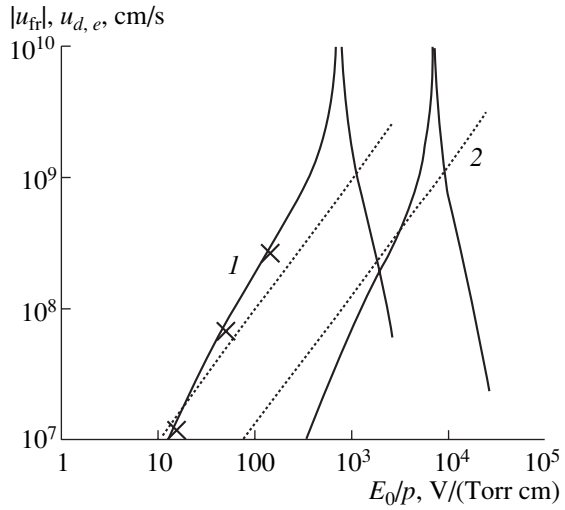


Fig. 4. Absolute value of the ionization front velocity $|u_{fr}|$ (solid curves) and the drift velocity $u_{d,e}$ (dotted curves) vs. reduced electric field at the streamer head E_0/p for (*1*) helium and (*2*) xenon at $N_{cr} = 10^{14}$ cm^{-3}, $N_0 = 10$ cm^{-3}, $r_0 = 0.5$ mm, and $p = 1$ atm (calculations by formulas (6)–(9)). The crosses show the computation results by model (1). Throughout the entire parameter range, $\alpha_i r_0 > 10$.

approximation was used [18, 20]:

$$\xi(x) = 45(\text{torr})^{-1} u_{d,e} \exp(-31.1(1/x)^{1/2} - 1.7\times 10^{-4} x), \quad (13)$$

$$u_{d,e} = \frac{1.3x + 1.3x^6}{1 + 7.31\times 10^{10} x^{5.8}} + 1.3\times 10^5 \exp\left(-\frac{2.2}{x}\right) \text{(cm/s)}. \quad (14)$$

For xenon, we have $(E_0/p)_{cr} \approx 7$ kV/(cm Torr). The front propagation velocity as a function of the reduced electric field in helium and xenon is shown in Fig. 4.

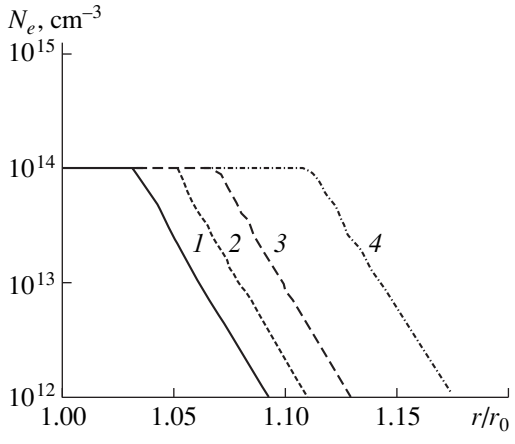


Fig. 5. Propagation of an ionization wave in helium in a nonuniform electric field for the same parameters as in Fig. 4 and for $E_0/p = 15.3$ V/(Torr cm): the electron density profiles at the times t = (*1*) 1.4, (*2*) 1.47, (*3*) 1.55, and (*4*) 1.75 ns (calculations by formula (1)).

Formula (12) was verified by numerically simulating a spherically symmetric plasma bunch. The electron density distributions at different times were calculated by formula (11) (see Fig. 5). Using these distributions, the time evolution of the front radius r_{fr} was approximated by a linear dependence (Fig. 6), which determines the front velocity. Some points thus obtained are shown in Fig. 4.

4. Front Velocity in N_2 and SF_6

When analyzing the propagation velocity of background multiplication in N_2 and SF_6 (see Fig. 7), we used the values of α_i and $u_{d,e}$ from Fig. 1 [19]. The nonmonotonic behavior of the propagation velocity of background multiplication in SF_6 is explained by the nonmonotonic dependence of the derivative of the ionization frequency with respect to the reduced field. This is caused by the presence of three threshold ionization energies (20, 40, and 50 eV) in SF_6.

COMPARISON WITH THE COMPUTATION RESULTS BY THE DRIFT–DIFFUSION MODEL

1. Drift–Diffusion Model

To confirm the presence of a background multiplication wave, we performed numerical simulations with the use of a one-dimensional drift–diffusion model (see [21] for details). This model accounts for the development of ionization between coaxial cylindrical electrodes, $r_0 < r < r_1$, where r_0 and r_1 are the radii of the inner and outer electrodes, respectively. The plasma formation and the electric field shielding were described by the equations for the momentum transfer and the continuity equations for electrons and ions, as well as by Poisson's equation for the electric field. The field dependences of various quantities entering into the

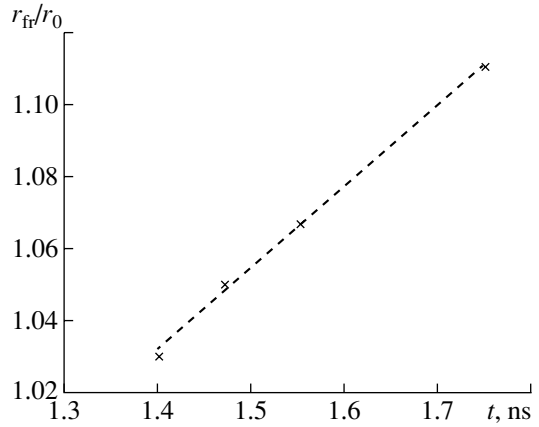


Fig. 6. Time evolution of the normalized front radius r_{fr}/r_0 for $E_0/p = 15.3$ V/(Torr cm). The crosses correspond to the data from Fig. 2. The dotted line shows the approximation $r_{fr}/r_0 = u_{fr}t/r_0 + 0.718$, with $u_{fr} = 1.1\times 10^7$ cm/s.

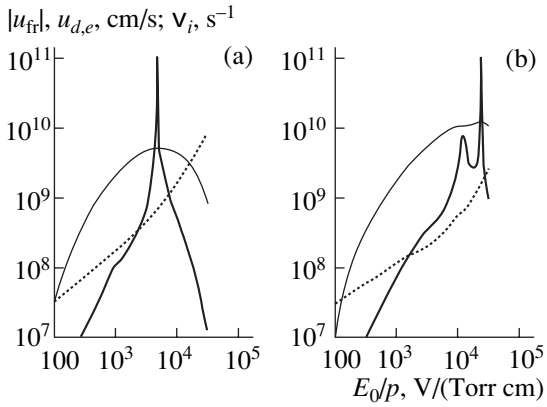


Fig. 7. Absolute value of the ionization front velocity u_{fr} (heavy solid curves), drift velocity $u_{d,e}$ (dotted curves), and ionization frequency ν_i (light solid curves) vs. reduced electric field E_0/p at the streamer head in (a) N_2 and (b) SF_6 for $N_{cr} = 10^{14}$ cm^{-3}, $N_0 = 10$ cm^{-3}, $r_0 = 0.5$ mm, and $p = 1$ atm; $(E_0/p)_{cr} \approx 4.7$ kV/(Torr cm) for N_2 and $(E_0/p)_{cr} \approx 23$ kV/(Torr cm) for SF_6. The data from [20, 21] were used.

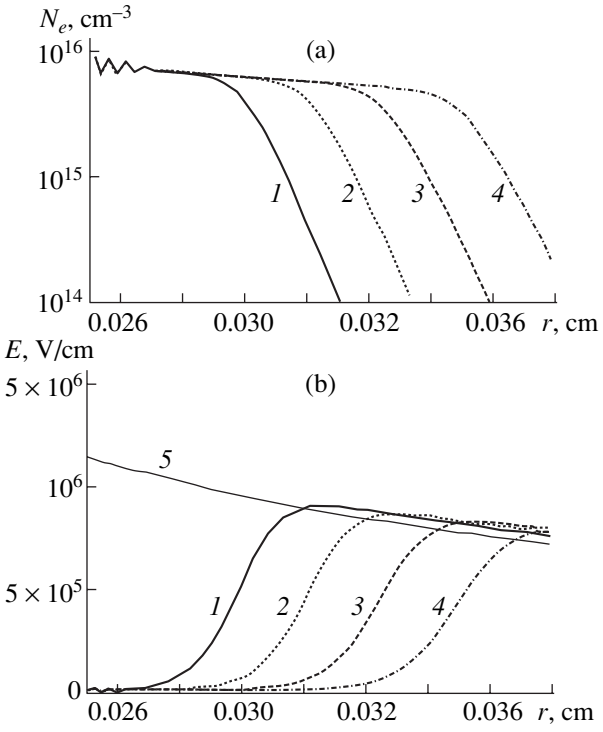


Fig. 8. Time evolution of the radial profiles of the (a) electron density and (b) electric field for a small-radius anode (the case of a cathode-directed background multiplication wave in xenon at $p = 1$ atm, $r_0 = 0.025$ cm, $r_1 = 0.825$ cm, and $U = 100$ kV): t = (*1*) 0.012, (*2*) 0.013, (*3*) 0.014, and (*4*) 0.015 ns (calculations by the comprehensive drift–diffusion model [20]). Curve *5* in Fig. 8b shows the electric field profile in vacuum.

equations of the drift–diffusion model (the ionization frequencies, drift velocities, and diffusion coefficients) were specified using the approximations for xenon from [20]. The voltage between the electrodes was specified, and the possibility was also ensured of specifying the initial uniform background plasma density and a region with an increased degree of ionization in the form of a Gaussian distribution.

Because of the very large amount of the data obtained, let us restrict our consideration to the brief description of some of them.

2. Plane Geometry

In the case of nearly plane electrodes ($d = r_1 - r_0 \ll r_1$), the ionization wave propagated from the cathode toward the anode only if an excess degree of ionization was set near the cathode at the initial time. Moreover, the anode-directed ionization wave was observed only at low voltages and, accordingly, small Townsend multiplication coefficients $\alpha_i \lesssim 1/d$, i.e., when the electrons escaped from the discharge gap without significant multiplication. In the opposite case (when $\alpha_i d \gg 1$), volume ionization proceeded faster than the electrons drifted toward the anode; thus, the region with an increased initial degree of ionization did not have time to significantly shift over the time required for the ionization of the entire gas in the gap.

Indeed, the electrons pass through the discharge gap over a time of $d/u_{d,e}$. The region with an elevated initial plasma density begins to expand when the characteristic time of volume ionization $Ln(\nu_i)^{-1} = Ln(\alpha_i u_{d,e})^{-1}$ is longer than or on the order of the electron drift time, $Ln/\alpha_i u_{d,e} < d/u_{d,e}$, i.e., when $\alpha_i d < Ln$. If there is no region with an elevated initial plasma density, then, according to Eq. (10), the velocity of the ionization front at $r_0 \longrightarrow \infty$ is infinitely high and ionization occurs simultaneously throughout the entire volume.

3. Nonuniform Electric Field

For coaxial cylindrical electrodes with a low-radius cathode ($r_1 \gg r_0$), a region with an elevated degree of ionization is formed near the cathode at both low and high voltages. This region propagates in accordance to expression (6) due to the electric field nonuniformity, rather than due to the electron drift.

To demonstrate the effect of a background multiplication wave, the computations were performed with a low-radius anode (i.e., a positive potential was applied to the electrode of radius r_0, while a zero potential was applied to the outer electrode of radius r_1). According to the computation results, the background multiplication wave propagates from the anode toward the cathode (Fig. 8). The propagation velocities calculated using the complete drift–diffusion model and background multiplication model (6) differ by only 3% (see Fig. 9).

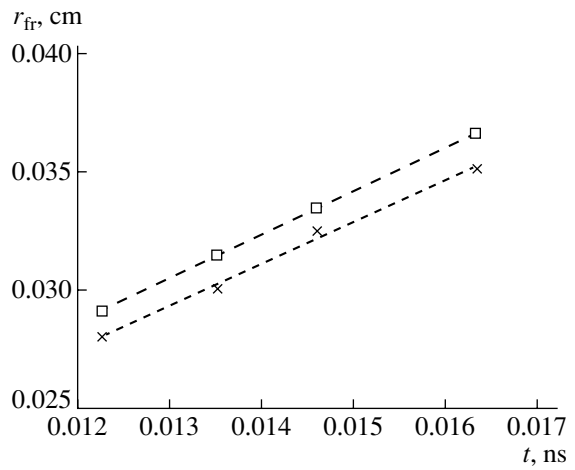


Fig. 9. Time evolution of the front radius r_{fr}. The crosses show the results of calculations by the comprehensive drift–diffusion model at the same times as in Fig. 4. The squares show the results of calculations by the background multiplication model (1). The dotted and the dashed lines show the linear approximations of the computation results: $r_{fr} = u_{fr}t +$ const, with $u_{fr} = 1.85 \times 10^8$ cm/s and const $= 6.4 \times 10^{-3}$ cm for the drift–diffusion model and $u_{fr} = 1.79 \times 10^8$ cm/s and const $= 6.06 \times 10^{-3}$ cm for the background multiplication model (1).

DISCUSSION

1. Streamer Propagation Velocity

When $E_0/p < (E_0/p)_{cr}$, the velocity of the background multiplication wave is on the order of the experimentally observed streamer velocity (~10^7–10^9 cm/s). A more precise comparison with the experimental data calls for two-dimensional computations. Note that, at certain values of r_0 and E_0/p, the velocity of the multiplication wave is equal to the electron drift velocity and exceeds it at higher electric fields and large head radii (see Figs. 4, 7). Hence, at low electric fields, the conventional mechanism for the propagation of the anode end of a streamer due to the electron drift can prevail in the initial stage. However, as the degree of ionization increases, the electric field at the streamer head increases (due to an increase in the streamer length). Therefore, in the later stage, the ionization due to background multiplication should be dominant.

As for the propagation of the cathode end of a streamer, it obviously cannot be caused by the electron drift. The propagation of the ionization wave toward the cathode is apparently related to the above-discussed background electron multiplication in a nonuniform electric field.

2. Extension of the Ionization Region

To calculate the streamer shape, it is necessary to use two- and even three-dimensional models. Some conclusions can, however, be made based on the above one-dimensional consideration. The shape of the streamer head is determined by the competition between two factors. According to Eq. (11), the front velocity increases with electric field (if $E_0/p < (E_0/p)_{cr}$) and increases linearly with the radius of curvature r_0. However, the electric field is higher at points where the radius of curvature is smaller. At moderate electric fields ($E_0/p < (E_0/p)_{cr}$), the dependence on the electric field is significantly steeper than linear. Therefore, the very end of the head will propagate faster than its other parts. Thus, at $E_0/p < (E_0/p)_{cr}$, the ionization region will stretch out.

It should be noted, however, that the minimum radius of curvature r_0 cannot be made arbitrarily small because the above consideration is valid for $\alpha_i r_0 \gg 1$.

The electric field at the streamer head increases with the streamer length. Therefore, the propagation velocity of a streamer should also increase with its length, as was experimentally observed in [6].

3. Infinite and Negative Velocities

It follows from Eq. (7) that, when $E_0/p = (E_0/p)_{cr}$, the propagation velocity of the background multiplication wave tends to infinity. Since the multiplication wave does not transfer information from point to point in the propagation direction, this velocity can formally be even higher than the speed of light. However, in some cases, the delayed shielding of the external field (not taken into consideration in the above model) should limit the propagation velocity of the multiplication wave to the speed of light.

When $E_0/p > (E_0/p)_{cr}$, the front velocity of the background multiplication wave is opposite to the field gradient. Hence, at a certain distance from the front (at the point where $E_0/p = (E_0/p)_{cr}$), an ionization region that extends in both directions should emerge. In this case, one should observe a jerky motion of the streamer. A slightly nonmonotonic dependence of the ionization frequency on the electric field (Fig. 7b) should also lead to an unstable streamer propagation.

CONCLUSIONS

Thus, a simple mechanism for the propagation of an ionization wave in a dense gas has been proposed. The mechanism does not depend on the sign of the projection of the electric field on the streamer propagation direction and is related to the difference in the rates of electron multiplication in a nonuniform electric field. The streamer propagation is caused by the enhancement of the electric field at the streamer head.

The multiplication of electrons in both electropositive and electronegative gases at prebreakdown fields has been considered. The initial ionization is enabled by the natural background of ionizing radiation and cosmic rays. It is shown that prebreakdown multiplication is capable of providing a sufficiently high density

of background electrons. This allows one to treat the background as a continuous medium when considering streamer propagation as a multiplication wave.

An analytic expression for the velocity of the ionization front is obtained based on a simple equation for the multiplication of background electrons. This expression is in good agreement with numerical simulations performed within both the simple model of background electron multiplication and a more comprehensive drift–diffusion model. In particular, it is shown using the drift–diffusion model that the ionization front can propagate from the small-radius anode to the cathode due to the multiplication of background electrons. The propagation velocity of the ionization wave is computed as a function of the electric field at the streamer head for He, Xe, N_2, and SF_6.

Obviously, in this paper, we did not take into consideration a great number of various phenomena related to the streamer propagation. In particular, preionization can be provided by the fast electrons emitted from the anode-directed end of the streamer. Also, it was not taken into account that a streamer can branch and that preionization in lightning can be provided by the fast (precursor) electrons. Nevertheless, even the above simple analysis shows that the mechanism proposed can be considered as a good alternative for the photon hypothesis.

REFERENCES

1. L. B. Loeb, *Fundamental Processes of Electrical Discharges in Gases* (Wiley, New York, 1939; Gostekhizdat, Moscow, 1950).
2. V. L. Granovskiĭ, *Electric Current in Gases* (Gostekhizdat, Moscow, 1952), Vol. 1 [in Russian].
3. P. A. Kaptsov, *Electronics* (Gostekhizdat, Moscow, 1954) [in Russian].
4. J. M. Meek and J. D. Craggs, *Electrical Breakdown of Gases* (Clarendon Press, Oxford, 1953; Inostrannaya Literatura, Moscow, 1960).
5. H. Raether, *Electron Avalanches and Breakdown in Gases* (Butterworths, London, 1964; Mir, Moscow, 1968).
6. É. D. Lozanskiĭ and O. B. Firsov, *Theory of Spark* (Atomizdat, Moscow, 1975) [in Russian].
7. Yu. P. Raizer, *Physics of Gas Discharge* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).
8. M. C. Wang and E. E. Kunhard, Phys. Rev. A **42**, 2366 (1990).
9. P. A. Vitello, B. M. Penetrante, and J. N. Bardsley, Phys. Rev. E **49**, 5574 (1994).
10. A. A. Kulikovsky, Phys. Rev. E **57**, 7066 (1998).
11. N. L. Aleksandrov, E. M. Bazelyan, *et al.*, J. Phys. D **32**, 2636 (1999).
12. A. Rocco, U. Ebert, and W. Hundsdorfer, Phys. Rev. E **66**, 035102(R) (2002).
13. M. Arrayás, U. Ebert, and W. Hundsdorfer, Phys. Rev. Lett. **88**, 174502 (2002).
14. A. A. Kulikovsky, Phys. Rev. Lett. **89**, 22940 (2002).
15. N. S. Rudenko and V. I. Smetanin, Izv. Vyssh. Uchebn. Zaved. Fiz., No. 7, 34 (1977).
16. L. P. Babich, Fiz. Plazmy **7**, 1419 (1981) [Sov. J. Plasma Phys. **7**, 783 (1981)].
17. A. N. Tkachev and S. I. Yakovlenko, Pis'ma Zh. Éksp. Teor. Fiz. **77**, 264 (2003) [JETP Lett. **77**, 221 (2003)].
18. A. N. Tkachev and S. I. Yakovlenko, Pis'ma Zh. Tekh. Fiz. **29** (16), 54 (2003) [Tech. Phys. Lett. **29**, 683 (2003)].
19. A. M. Boĭchenko, A. N. Tkachev, and S. I. Yakovlenko, Pis'ma Zh. Éksp. Teor. Fiz. **78**, 1223 (2003) [JETP Lett. **78**, 709 (2003)].
20. A. N. Tkachev and S. I. Yakovlenko, Laser Phys. **12**, 1022 (2002).
21. A. N. Tkachev and S. I. Yakovlenko, Laser Phys. **13**, 1345 (2003).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 49, No. 9, 2004, pp. 1158–1165. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 55–62.

SOLIDS

Equilibrium States of Planar Vortices in a Three-Dimensional Josephson Medium and Meaning of the Pinning-Energy Concept

M. A. Zelikman

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

Received December 8, 2003

Abstract—Various ways of specifying the pinning-energy concept for planar vortices in a three-dimensional cellular Josephson medium are analyzed. It is shown that, for values of the pinning parameter I that are not small, a universal characteristic of vortex interaction with the lattice cannot be found, since the displacement of a vortex distorts its shape. At small values of I, the maximum pinning force can be chosen for such a characteristic. Two equilibrium states of a vortex are analyzed for stability. It is revealed that the state of higher energy is not inevitably unstable. A correct analysis of stability must be based on exploring a quadratic form that describes the energy of a current configuration. Such an investigation is performed for the equilibrium state of a vortex. At small values of the pinning parameter, the vortex state of higher energy is quasistable. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Vortex structures play an extremely important role in processes occurring in superconducting materials. For this reason, a large number of theoretical and experimental studies have been devoted to examining the dynamics of vortices, their interaction with one another, their stability, their pinning on various inhomogeneities of a medium, and other relevant phenomena. The strength of pinning is usually characterized by the pinning energy. However, the meaning of this concept calls for refining. It is the most convenient to consider this issue within a model that admits an exact mathematical treatment.

Planar (laminar) vortices in a three-dimensional ordered Josephson medium that has the form of a cubic lattice consisting of superconducting wires and containing one Josephson contact in each connection were considered in [1].

The energy of a stationary vortex per 1 m of its length is given by

$$E = E_0 \sum_{k=-\infty}^{\infty} \left[\frac{1}{2}(\varphi_{k+1} - \varphi_k)^2 + I(1 - \cos\varphi_k)\right], \quad (1)$$

where φ_k are the phase jumps across the contact, I is the so-called pinning parameter, and E_0 is a normalization constant.

In the following, energies are everywhere expressed in terms of E_0 units. In expression (1), the first term represents the magnetic energy of a vortex, while the second term describes the energy of Josephson contacts.

The equilibrium states of a vortex, which correspond to an extremum of its energy, are described by the set of finite-difference equations

$$\varphi_{m+1} - 2\varphi_m + \varphi_{m-1} = I\sin\varphi_m, \quad (2)$$

which are obtained from the condition $\partial E/\partial\varphi_m = 0$.

A set of equations that is similar to (2) also describes the behavior of vortices in a long periodically modulated Josephson contact [2, 3].

The set of Eqs. (2) was first obtained by Frenkel and Kontorova in describing the behavior of edge dislocations in crystals [4]. In view of its discreteness, this set of equations makes it possible to analyze the behavior of solitons with allowance of their pinning, which is controlled by the energy that is required for shifting the center of a vortex from one cell to another.

For the set of Eqs. (2), solutions of two types corresponding to equilibrium planar vortices were considered in [1]. In one of these, the symmetry axis passes through the center of a medium cell. In the other, the center of a vortex lies on one of the wires. Since the energy of the second configuration exceeded the energy of the first configuration, it was concluded in [1] that the first configuration is stable, but that the second configuration is unstable. In addition, it was tacitly assumed in [1–3] that a vortex that initially has the first configuration and which moves to the neighboring cell traverses half a cell in the coordinate and takes the form of the second configuration, which is the configuration that corresponds to the maximum energy of the vortex at the instant of its passage above the barrier separating the cells. On the basis of this model, it was deduced that the pinning energy is precisely the energy difference between the second and the first configuration of a vortex.

In the present study, it is shown that both these conclusions—that concerning the character of stability and

that concerning the mode of vortex motion—are incorrect and that the concept of the pinning energy requires refining.

The character of stability is not determined unambiguously by the hierarchy of energies of equilibrium states. The configuration having the lowest energy is always stable. In general, the second configuration can also be stable, despite the fact that its energy exceeds the energy of the first configuration. A configuration of higher energy must be individually analyzed for stability. Prior to clarifying the character of stability, we will not therefore apply the terms "stable" and "unstable" to the aforementioned configurations—instead, we will merely refer to them as the first and the second configuration.

At the instant of a transition from one cell to another, a vortex is not in its second equilibrium state; therefore, the barrier energy is not equal to the energy difference between the second and the first configuration. In order to prove this statement, it is convenient to consider the behavior of the system formed by two isolated vortices.

TWO ISOLATED INTERACTING VORTICES

Let us analyze the set of Eqs. (2) for the case of two vortices that have identical orientations [5], which are far off the boundary, and whose centers separated by N rows occur in rows 1 and $(N + 1)$ (the numbers of the rows are enclosed by circles in Fig. 1). As one moves to the left of cell 1 and to the right of cell $(N + 1)$, the currents J_k decrease, tending to zero at infinity.

At values of the pinning parameter I that are not small, the values of φ_m off the central cells of the vortices are small, the set of Eqs. (2) linearizes, and its solution has the form [5]

$$\varphi_m = \varphi_{-1}\gamma^{-m-1} \quad (m \le -1), \tag{3}$$

$$\varphi_m = C_1\gamma^{m-1} + C_2\gamma^{N-m} \quad (1 \le m \le N), \tag{4}$$

$$\varphi_m = \varphi_{N+1}\gamma^{m-N-1} \quad (m \ge N+1), \tag{5}$$

where $\gamma = 1 + I/2 - \sqrt{I + I^2/4}$ is a solution to the characteristic equation $\gamma^2 - (2 + I)\gamma + 1 = 0$.

We will assume that the distribution of the currents and phases is symmetric with respect to the midpoint of the distance between the vortices. For identically oriented vortices that are repelled from each other, we then have $C_1 = -C_2 = \varphi_1/(1 - \gamma^{N-1})$ and

$$\varphi_2 = \varphi_1 k, \tag{6}$$

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

Substituting (6) into the boundary conditions in the central cell of the vortex,

$$I\sin\varphi_1 = \varphi_{-1} - 2\varphi_1 + \varphi_2 + 2\pi, \tag{7}$$

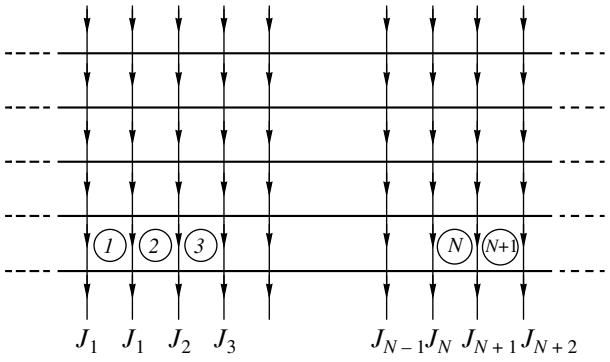


Fig. 1. Distribution of currents in a plane orthogonal to the vortex axes for two solitary interacting vortices.

$$I\sin\varphi_{-1} = \varphi_1 - 2\varphi_{-1} + \varphi_{-2} - 2\pi, \tag{8}$$

we obtain the following set of equations for determining φ_{-1} and φ_1:

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

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

In Eqs. (9) and (10), only the coefficient k depends on the distance between the vortices: $k(\infty) = \gamma$, $k(2) = -1$, $k(3) = 0$, and so on. In Fig. 2, the graphs corresponding to Eqs. (9) and (10) are given for $I = 2.0$ at $N = 2$, 4, and ∞ and for $I = 2.9$ at $N = 3$ and ∞. On the basis of these curves, it is concluded in [5] that, at each value of I, the set of Eqs. (9) and (10) has a solution in a specific range of N values—that is, vortices can be separated by different distances from a minimum distance to infinity. This is the result of pinning, since, in the absence of pinning, the vortices would go to infinity—that is, there would be no solutions at finite N values. By determining the minimum value of N at given I, we will find the minimum distance at which pinning forces can compensate the mutual repulsion of the vortices. From Fig. 2a, it can be seen that, at $I = 2$, $N_{\min} = 2$. If the centers of the vortices lie in neighboring cells (that is, if $N = 1$), it follows from (6) that $k(1) = \infty$; from (9), we then obtain $\varphi_1 = 0$, which is readily understandable in view of the symmetry of the picture. The minimum value of I at which Eq. (10) is still satisfied at $\varphi_1 = 0$ is 2.9. In Fig. 2b, the curve that is tangent to the φ_{-1} axis at the point of minimum corresponds to this value.

In a similar way, we can also consider the case of vortices having opposite orientations and attracting each other. In that case, we set $C_2 = C_1$ in (4) and arrive at Eqs. (9) and (10), the only difference being that, here, $k = \gamma(1 + \gamma^{N-3})/(1 + \gamma^{N-1})$. Figure 3 shows the graphs for this case at $I = 2.16$ and $N = 3$, $I = 3.69$ and $N = 2$, and $I = 14.3$ and $N = 1$.

On the basis of the data in Figs. 2 and 3, we can draw yet another important conclusion. The graphs corresponding to Eqs. (9) and (10) have three or even five intersection points rather than two, as might have been

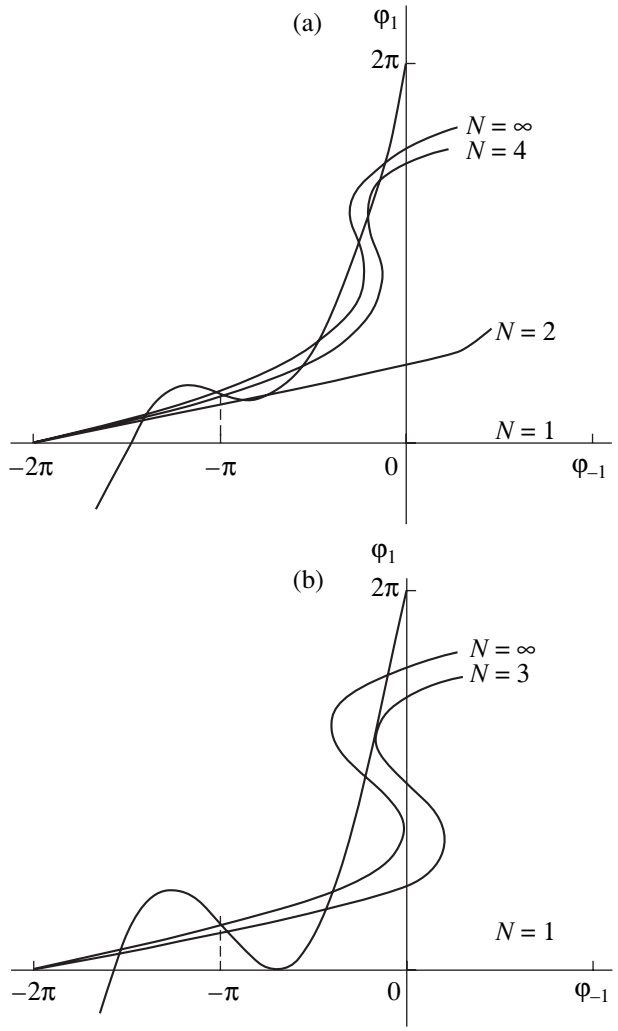


Fig. 2. Graphical solution to the set of Eqs. (9) and (10) for vortices of identical orientations (repulsion) at various values of the pinning parameter I: (a) 2.0 and (b) 2.9.

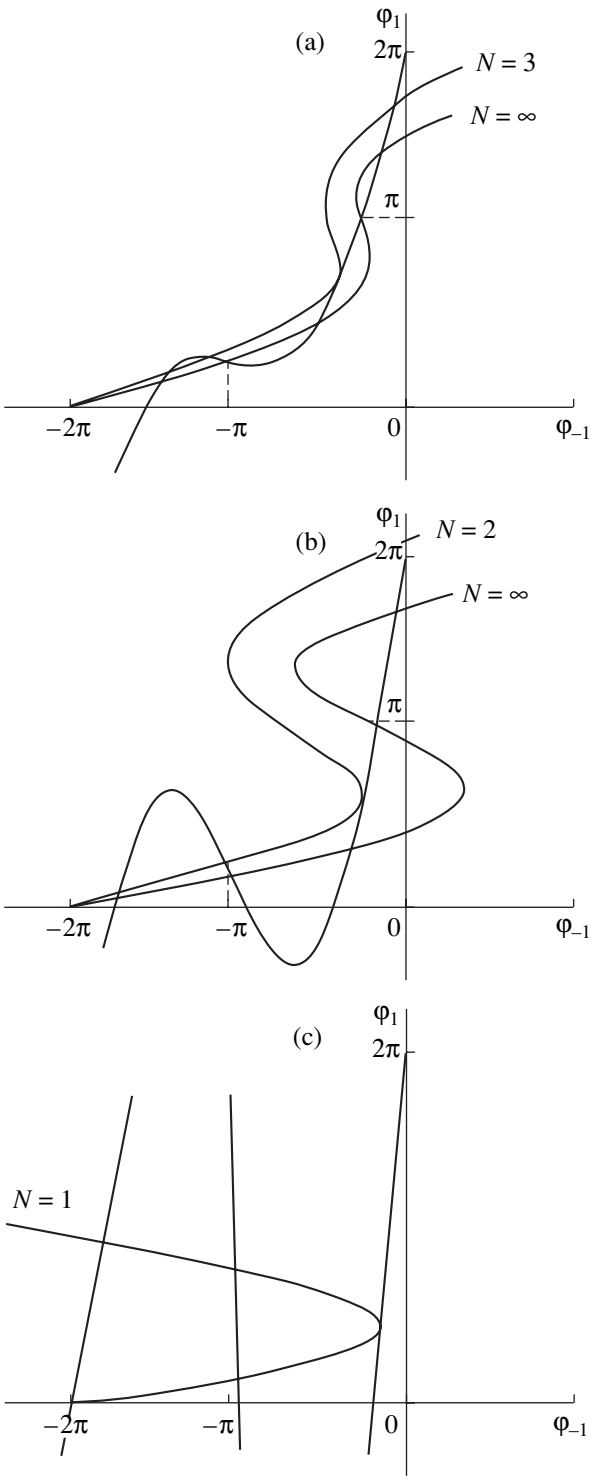


Fig. 3. Graphical solution to the set of Eqs. (9) and (10) for vortices of opposite orientations (attraction) at various values of the pinning parameter I: (a) 2.16, (b) 3.69, and (c) 14.3.

expected in view of the fact that a vortex has two equilibrium states. This, however, does not indicate that there are additional equilibrium vortex configurations. As a matter of fact, there are only two intersection points, the remaining ones being their copies that correspond to the subtraction of 2π in a different row of contacts (this does not change the physical situation).

Let us consider in greater detail this statement for the case of a solitary vortex, which corresponds to $N = \infty$. We take the central intersection point at $-\varphi_{-1} = \varphi_1 = a$. For the vortex to look symmetric in the coordinate, 2π was added in this case to the results obtained on the basis of (2) in calculating φ_1. If this had been done in calculating φ_2, then φ_1 would have been equal to $(a - 2\pi)$—that is, it would have been negative and would have been considered as a new phase φ_{-1}. In this case, the former phase $\varphi_2 = a\gamma$ [from (4)] would have become a new phase φ_1. These are the values that correspond to the extreme left point of the intersection of the curves associated with Eqs. (9) and (10). Following a similar line of reasoning, we find that, if the term 2π had been added in calculating φ_{-1} rather than φ_1, then the former

phase $\varphi_{-2} = -a\gamma$ would have become a new phase φ_{-1}, while the former phase φ_{-1} transformed into $(-a + 2\pi)$ would have been a new phase φ_1. This point corresponds to the extreme right point of the intersection of the curves in Fig. 2. Thus, both of the extreme intersections correspond to a physical state that is described by the central intersection point—that is, to the first vortex configuration. In a similar way, one can show that the remaining two intersection points (for $\varphi_{-1} = -\pi$ and $\varphi_1 = \pi$) both correspond to the second configuration.

For the curves associated with $N \neq \infty$—that is, for the case of two interacting vortices—the situation is similar, but physically identical intersection points lie on different curves. Indeed, we recall that, since, in deriving Eqs. (9) and (10) for repelling vortices, we assumed the distribution of the phases to be symmetric (that is, $C_1 = -C_2$), the disposition of a new phase φ_{-1} in terms of the former phase φ_1 corresponds to a decrease of 2 in the distance N between the vortices. As identical intersection points, we then have, for example, the central intersection point at $N = 4$ and the extreme left intersection point at $N = 2$, the extreme right intersection point at $N = 4$ and the central intersection point at $N = 2$, and the second intersection point from the right at $N = 4$ and the second intersection point from the left at $N = 2$. This can clearly be seen in Fig. 2a, where the curves specified by Eqs. (9) and (10) are tangent to each other at $N = 4$ on the upper right and at $N = 2$ on the lower left, this confirming the identity of two pair of points simultaneously. A similar pattern can be seen in Fig. 2b, where, at $N = 3$, the curves touch each other on the upper right and where, at $N = 1$, the curve specified by Eq. (10) touches the abscissa, which, at $N = 1$, corresponds to (9).

Thus, there are only two physically different states. This conclusion is valid for the case of two attracting vortices as well (see Fig. 3). Further, states corresponding to the central intersection point in Fig. 2 and the second intersection point from the left ($\varphi_{-2} = -\pi$ for $N = \infty$) will be considered as the representative states for repelling vortices. In the case of attracting vortices, the curve specified by Eq. (9) deviates in the opposite direction as the vortices approach; therefore, it is more convenient to consider the central intersection and the second one from the right ($\varphi_1 = \pi$ for $N = \infty$) as the representative states in this case.

ENERGY BARRIER

The above analysis of two interacting vortices presumes a jumplike change in the distance N between the centers of the vortices. This means that, at some value of N, points at which the curves in question intersect in Figs. 2 and 3 still exist, but that there are no such points at N less than that by unity. Therefore, this approach gives no way to consider the exact instant of the transition from one cell to another, this instant corresponding to the state of maximum energy.

Instead of a symmetric configuration of two vortices, we will now consider one asymmetric vortex characterized by the following values of φ_k across the contacts closest to the center:

$$\varphi_{-1} = -a, \quad \varphi_{-2} = -a\gamma, \quad \varphi_1 = b - d\gamma, \quad \varphi_2 = b\gamma - d. \tag{11}$$

These values correspond to expressions (3) and (4). Substituting (11) into Eqs. (7) and (8), we arrive at

$$\varphi_{-1} = I\sin\varphi_1 + (2-\gamma)\varphi_1 - 2\pi + d(1-\gamma^2), \tag{12}$$

$$\varphi_1 = I\sin\varphi_{-1} + (2-\gamma)\varphi_{-1} + 2\pi. \tag{13}$$

Here, positive (negative) values of d correspond to repulsion (attraction) on the right, the growth of d leading to the enhancement of this interaction, with the result that, at some value of d, the vortex being considered can no longer occur in a given cell and must jump into the next one. This is the value of d that corresponds to the energy barrier.

The set of Eqs. (12) and (13) differs from the set of Eqs. (9) and (10) only by the presence of the term $d(1 - \gamma^2)$ in (12). This term shifts the entire curve $\varphi_{-1}(\varphi_1)$ at $k = \gamma$ to the right (or to the left). Figure 4 shows the graphs specified by Eqs. (12) and (13) at $I = 2.0$ and some values of d for the cases of (curve *1*) attraction and (curve *2*) repulsion. It was shown above that there are only two physically different configurations. The corresponding points of the intersection of the graphs associated with Eqs. (12) and (13) are represented by the closed circles in Fig. 4. One can see that, with increasing d, these points approach each other, merging at some value of d, whereupon there are no representative intersection points. This means that the vortex can no longer remain in the cell being considered. It is

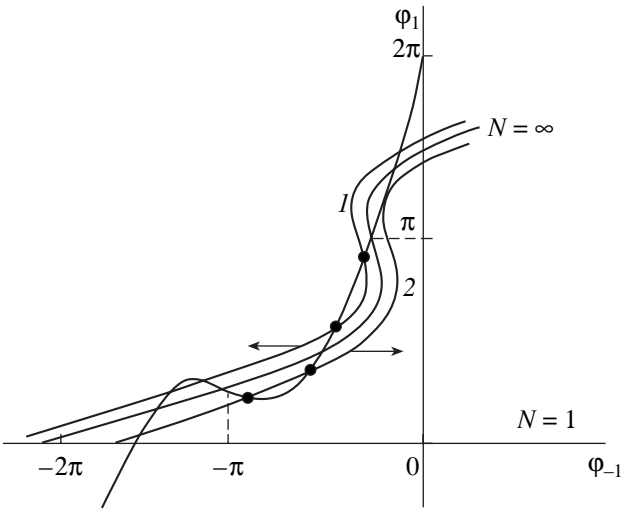


Fig. 4. Graphical solution to the set of Eqs. (12) and (13) at $I = 2.0$ and some values of d for the cases of (*1*) attraction and (*2*) repulsion.

Energies E_b corresponding to the jump of a vortex to a different cell for the cases of repulsion and attraction and energies E_1 and E_2 of the equilibrium states of a solitary vortex

No.	Repulsion				Attraction			
	I_{min}	E_b	E_1	E_2	I_{min}	E_b	E_1	E_2
1	2.907	14.85	12.36	13.07	14.30	22.31	17.43	37.86
2	1.947	11.46	10.62	10.99	3.690	15.78	13.36	15.42
3	1.428	9.65	9.28	9.40	2.156	14.13	11.07	11.57

important to note that, in contrast to what was implied in [1–3], this energy-barrier state is not the second configuration of an equilibrium vortex. Moreover, a vortex does not assume its equilibrium configurations (neither the first nor the second one) at any instant while going over from one cell to another. Even after the transition to the next cell, d is not equal to zero—that is, the vortex does not assume its first equilibrium configuration.

The aforementioned assumptions were based on the erroneous analogy between a vortex and a pointlike object on a potential relief. Indeed, the energy barrier for a pointlike object going over from one valley to another corresponds to the position of the point on the ridge—that is, to its unstable equilibrium. Upon overcoming the barrier, the point occurs in a state of a stable equilibrium.

A vortex is not a pointlike object. It is a configuration whose shape depends on the situation.

Let us consider the energy in (1) as a function of the variables φ_k. This function describes a "mountain system" in a multidimensional state. The equilibrium states described by the set of Eqs. (2) correspond to the vanishing of all partial derivatives; therefore, they can be points of a local maximum or a minimum or saddle points. Each of these points corresponds to an equilibrium state, these states involving various numbers of vortices occurring at various distances from one another. It is quite obvious that a configuration involving several vortices has, even in the case of their mutual attraction, a total energy that exceeds the energy of a solitary vortex. This means that, among all critical points of the system (that is, equilibrium states), one can single out states that possess the lowest energies and which contain one solitary vortex. The number of such points is indefinitely great, but all of them correspond to the situation where only two physically different configurations, referred to as the first and the second one, are arranged in various cells of the medium. Obviously, the first configuration, which has a lower energy, corresponds to a minimum, since there cannot be a lower energy at all. Any infinitesimal change in the combination φ_k leads to an increase in energy; therefore, this equilibrium state is stable. In contrast to what was tacitly assumed in [1–3], the second configuration does not correspond to a maximum, albeit it has a higher energy. In fact, this point is surrounded by other states corresponding to several vortices and having an energy higher than its energy. Thus, the second configuration may correspond either to a minimum or to a saddle point. This state would be stable in the former and unstable in the latter case, this being so with respect to some specific combinations of fluctuations rather than with respect to all of them. A dedicated investigation is necessary for revealing the character of stability of the second configuration, and it will be performed below.

In analyzing the motion of a vortex from one cell to another, we cannot consider a vortex as a solitary object since it changes its shape. We assume that the force causing the motion of a vortex is due to its interaction with another vortex. The energies E_b corresponding to the jump of a vortex to a different cell for the cases of repulsion and attraction are given in the table along with the energies E_1 and E_2 of two equilibrium states of a solitary vortex. The distances N between the centers of the vortices are presented in the first column, and the smallest values of the pinning parameter I at which these distances may take such values are quoted in the second and sixth columns. All of these values were obtained on the basis of a precise numerical solution to the set of Eqs. (2). The fact that the I_{min} values are much less in the case of repulsion than in the case of attraction is worthy of special note. This indicates that, at the same distance between the vortices, the force of attraction between oppositely oriented vortices is much greater than the force of repulsion between identically oriented vortices. In this respect, the situation is analogous to that in the electrostatic interaction of charged spheres having finite dimensions (that is, those that cannot be considered as pointlike charges): at identical distances between the spheres and identical absolute values of the charges, the force of attraction between unlikely charged spheres is greater than the force of repulsion between likely charged spheres.

ON THE MEANING OF THE PINNING-ENERGY CONCEPT

In the situation being considered, it would be reasonable to define the pinning energy as $E_b - E_1$. From the table, one can see that, in the case of repulsion, this quantity exceeds considerably the difference $E_2 - E_1$, which was taken for the pinning energy in [1–3]. In the case of attraction, $E_b - E_1$ may be either greater or less

than $E_2 - E_1$. Therefore, the difference $E_b - E_1$ is not a universal characteristic of pinning either. It depends on the mechanism that displaces a vortex—in particular, on whether the vortex being considered is attracted by its neighbor from one side or is repelled by its neighbor from the other side. Different mechanisms of action on a vortex may also exist. In each case, the quantity $E_b - E_1$ depends on the character of variation of the entire combination φ_k. In particular, there can arise the question of the minimum vortex energy at which the vortex being considered could move by inertia. The answer to this question is also ambiguous, because such a motion can be organized in different ways, depending on the initial combination of the differences of phases and on the rates of their variation, an individual energy value corresponding to each mode of the motion.

The following model is useful in analyzing the concepts of a pinning energy and a pinning force. For the sake of simplicity, we will treat vortices as interacting spheres each of which occurs in an individual well surrounded by a barrier. At the instant corresponding to the escape of a vortex from its well, the energy of this vortex is then equal to the sum of the energy of the solitary vortex, the energy of its interaction with a neighboring vortex, and the positive barrier energy. In the case of repulsion, the vortex-interaction energy is positive. Therefore, the total energy of the vortex is greater than the solitary-vortex energy. In the case of attraction, the interaction energy is negative, whence it follows for the second configuration that, for a large distance between the vortices ($N = 3$), $E_b > E_2$ since the absolute value of the interaction energy is less than the barrier energy, but, for $N = 1$, the hierarchy of the energies in question is opposite. For the first configuration, however, this argument does not work, since we always have $E_b > E_1$. In all probability, this is because a vortex cannot be considered as a sphere.

Within the model being considered, the well depth or the barrier height would be the most natural choice for the pinning energy, because these are precisely the quantities that characterize the interaction of a vortex with the "lattice." The problem to be solved here is that of separating the energies of vortex interactions with neighbors and with the lattice. The point is that the model that treats a vortex as a pointlike object (sphere) is simplified: it disregards changes in the shape of a vortex, which lead to a change in the energy of its interaction with other vortices and a change in the well depth, depending on the situation. Thus, this quantity cannot be a universal characteristic of pinning either.

For the same reason, it cannot be stated that a vortex will jump from one cell to another as soon as the force acting on it owing to the interaction with other vortices exceeds some universal value of the pinning force. According to its meaning, the former of these forces must be equal to the gradient of the energy of the interaction of this vortex with other vortices, while the latter is expected to be equal to the gradient of the energy of its interaction with the lattice. In different situations, however, a vortex has different shapes; therefore, the separation of these two energies and, hence, of the respective forces is questionable.

A consideration based on studying the set of equations describing a vortex for a specific problem is strictly speaking the only correct approach in this case. The strength of pinning can be characterized, for example, by the minimum possible distance between two identical isolated vortices, as was done above (see Figs. 2, 3).

The approach proposed above and the set of Eqs. (9) and (10) are valid under the assumption that the pinning parameter I is not small. For I values less than or on the order of unity, this approach is not legitimate.

At small values of I, pinning was analyzed in [1–3, 5]. In view of the aforesaid, the procedures employed there seem incorrect for the following reasons. In [1–3], the energy of a vortex for various displacements of its center with respect to the center of a cell was calculated by formula (1) for a vortex having an invariable shape and moving along a coordinate. However, the shape of a vortex is not invariable. For a vortex, any position (at which its center does not lie either at the center of a cell or on a wire) can be an equilibrium position if a force generated, for example, by other vortices acts on it. This means that the vortex shape can no longer be symmetric—one of its sides ascends (or descends in the case of repulsion). The total energy of a vortex must include the energy of its interaction with other vortices; therefore, it differs from the results obtained in [1–3], where attention was given primarily to mathematical aspects of a calculation of the vortex energy without taking into account deviations of its shape from a symmetric shape. This distinction is confirmed by the calculations performed in [5], where it was tacitly assumed that, at small values of I, the energies of the interaction of a vortex with other vortices and with the lattice, as well as the corresponding forces, can be calculated separately. The former is calculated without allowance for pinning on the basis of the model of continuous vortices, while the latter is calculated without allowance for interaction with other vortices—that is, for the symmetric case considered in [1–3]. On the basis of this approach, the minimum possible distances between planar vortices were calculated in [5] from the condition that the force generated by the interaction of neighboring vortices is equal to the maximum pinning force. It can easily be shown that, under this condition, the vortex-interaction energy is $2\pi/\sqrt{I}$ times greater than the energy difference $(E_2 - E_1)$ between the two equilibrium states of a vortex—that is, the energy of its interaction with the lattice. Thus, we see that, at the instant corresponding to the transition of a vortex from one cell to another, the total energy of its interaction exceeds the difference $(E_2 - E_1)$ considerably.

Despite these reservations, however, the values obtained in [5] for the distances between vortices are

(a)

$2 + I\cos\varphi_N$	…	0	0	0	0	0	…	0
…	…	…	…	…	…	…	…	…
0	…	$2 + I\cos\varphi_2$	–1	0	0	0	…	0
0	…	–1	$2 + I\cos\varphi_1$	–1	0	0	…	0
0	…	0	–1	2–I	–1	0	…	0
0	…	0	0	–1	$2 + I\cos\varphi_1$	–1		0
0	…	0	0	0	–1	$2 + I\cos\varphi_2$	…	0
…	…	…	…	…	…	…	…	…
0	…	0	0	0	0	0	…	$2 + I\cos\varphi_N$

(b)

a_N	…	0	0	0	0	0	…	0
…	…	…	…	…	…	…	…	…
0	…	a_2	–1	0	0	0	…	0
0	…	–1	a_1	–1	0	0	…	0
0	…	0	0	$2 - I - 2/a_1$	–1	0	…	0
0		0	0	0	a_1	–1	…	0
0	…	0	0	0	0	a_2	…	0
…	…	…	…	…	…	…	…	…
0	…	0	0	0	0	0	…	a_N

Fig. 5. Matrix of the quadratic form in relation (14) for the second configuration: (a) original matrix and (b) result of reducing it to a quasidiagonal form.

close to those that were obtained on the basis of a precise numerical solution to the set of Eqs. (2). This confirms the validity of the assumption that the interaction of a vortex with other vortices and its interaction with the lattice can be calculated separately at small I, whence we draw the following important conclusion: at small values of the parameter I, pinning can be characterized by a universal value of the maximum pinning force. It was shown above that, at I values that are not small, the force confining a vortex depends on how this vortex is displaced—that is, it is not a universal characteristic.

STABILITY OF THE EQUILIBRIUM STATES OF A VORTEX

As was shown above, a vortex is not a pointlike object. It is a configuration whose shape changes, depending on the situation. Therefore, one has to consider, instead of the stability of a vortex as a discrete unit with respect to a displacement (this would be similar to what we have in the case of a sphere on a potential relief), stability with respect to any fluctuations of the set of φ_k values—that is, with respect to any transformations of the shape.

We will consider the energy as a function of the set of variables φ_k. The combination of the variables φ_k that describes the first configuration corresponds to the minimum energy of a solitary vortex. This point of a multidimensional space occurs at a local minimum of the function in (1) and is therefore stable.

Let us investigate the second configuration for stability. For this, we write the increment of the energy in (1) for the set of variations $\delta\varphi_k$ [the first derivatives vanish by virtue of Eqs. (2)]. We have

$$\delta E = \sum_{k,m} \frac{\partial^2 E}{\partial\varphi_k \partial\varphi_m} \delta\varphi_k \delta\varphi_m. \quad (14)$$

Further, we assume that those φ_k that are rather far off the center of the vortex being considered ($k \geq N + 1$) are equal to zero. The matrix $(\partial^2 E/\partial\varphi_k\partial\varphi_m)$ of the quadratic form in (14) then becomes finite and, for the second configuration, assumes the form shown in Fig. 5a.

We add the first row multiplied by $1/(2 + I\cos\varphi_N)$ to the second row; further, we add the resulting second row multiplied by $1/(2 + I\cos\varphi_N)$ to the third row and proceed in this way up to the central row. Diagonalizing the matrix from the opposite angle, we add the ultimate column multiplied by $1/(2 + I\cos\varphi_{N-1})$ to the penultimate column and proceed in this way up to the central column. The resulting matrix, whose elements below the diagonal are all equal to zero, is given in Fig. 5b. The diagonal elements are calculated according to the recursion relations

$$a_N = 2 + I\cos\varphi_N, \quad a_{n-1} = 2 + I\cos\varphi_n - 1/a_n. \quad (15)$$

Upon performing calculations on the basis of (15) by using the values of φ_k for the second configuration that were obtained from a precise numerical computation in [5], it can be proven that, for all values of I and rather large values of N (vortex size), the inequality $a_1 < 2/(2 - I)$ always holds—that is, the central diagonal element $(2 - I - 2/a_1)$ is always negative. By way of example, we indicate that, at $I = 0.5$, $a_1 = 1.3375$ for $N = 3$, $a_1 = 1.3316$ for $N = 4$, and $a_1 = 1.3296$ for $N = 10$, while $2/(2 - I) = 1.3333$. At $I = 0.2$, $a_1 = 1.1192$ for $N = 5$, $a_1 = 1.1111098$ for $N = 15$, and $a_1 = 1.1111088$ for $N = 40$, while $2/(2 - I) = 1.1111111$. It can be seen that, at modest values of I, the calculation of a_k must be performed quite accurately. The fact that, at small values of N, the condition $a_1 < 2/(2 - I)$ does not hold any longer indicates that, at the value of I being considered, the vortex size is much larger than N, so that it is not legitimate to set all φ_k to zero for $k \geq N + 1$. In this case, N must be increased; at sufficiently large values of N, a_1 approaches saturation; that is, it ceases to change with increasing N. This is the value of a_1 that is true.

The determinant of a matrix whose elements below the diagonal are all zero is equal to the product of all diagonal elements. From the fact that at least one diag-

onal element is negative, it follows that the quadratic form in (14) is not positive definite; that is, the energy does not have a local minimum at this point. This means that the second configuration is strictly speaking unstable at any value of I and that the authors of [1–3] were right to treat it in this way.

However, the following argument must be taken into account. The reduction of the quadratic form in (14) to a strictly diagonal form will lead to $\delta E = \sum_i A_i \delta\xi_i^2$, where the quantities $\delta\xi_i$ are linear combinations of the original variations $\delta\varphi_k$ of the phase jumps. As a result, all of $\delta\xi_i$ will be different from zero at a nonzero value of the variation of any one of the phase jumps φ_k. In this case, a negative (but small in magnitude) value of one of the coefficients A_i will be unable to outweigh the effect of the other terms, which are positive, and to render δE negative. There is only a small probability that the set of variations $\delta\varphi_k$ will accidentally be such that all of the remaining variations $\delta\xi_i$ will be so small that they will not be able to compensate for a negative term. It is straightforward to understand that the smaller the negative coefficient in absolute value, the lower the probability of such an event. This means that the smaller the pinning parameter I, the less unstable the second configuration. At small values of I, we can say that this state is quasistable.

It should be noted that a similar analysis of the first configuration for stability confirms that this configuration is stable for all values of I. In this case, the central diagonal matrix element reduces to the form $(a_1 - 1/a_1)$, the calculation on the basis of the recursion relations (15) leading to $a_1 > 1$ for any value of I. This means that all diagonal elements are positive; that is, the quadratic form being considered is positive definite.

CONCLUSION

Two equilibrium states of planar vortices in a three-dimensional Josephson medium have been considered. The meaning of the concept of the pinning energy with respect to the transition of the vortex center from one cell to another has been examined. It has been shown that, in general, there are no grounds to take the energy difference between the equilibrium states for the pinning energy.

The possibility of considering, as the pinning energy, the difference of the energy of a vortex at the instant of its transition from one cell to another and the minimum energy of a solitary vortex has been analyzed. At values of the pinning parameter I that are not small, this quantity is not a universal characteristic of pinning either, since it depends on the mechanism displacing a vortex. In considering a specific situation, it should be borne in mind that the energy of a vortex is the sum of the energy of the solitary vortex, the energy of its interaction with other vortices, and the positive energy of its interaction with the lattice. Within this model, it would be the most natural to take, for the pinning energy, the depth of the well that characterizes the interaction of a vortex with the lattice. Here, the problem is that, at values of I that are not small, it is hardly possible to separate the energy of the interaction of a vortex with its neighbors and the energy of its interaction with the lattice, since the shape of a vortex changes as it moves. Therefore, the force confining a vortex is dependent on how it is displaced—that is, it is not a universal characteristic.

At small values of I, the energy of the interaction of a vortex with other vortices and the energy of its interaction with the lattice, as well as the respective forces, can be calculated separately: the calculations for the interaction with neighboring vortices disregard pinning and rely on the continuous-vortex model, while the calculations for the interaction with the lattice take no account of the interaction with other vortices. It follows that, at small values of the parameter I, pinning can be characterized by the universal value of the maximum pinning force.

Two equilibrium states of a vortex have been analyzed for stability. It has been shown that the state of higher energy is not inevitably unstable. A correct analysis of stability must be based on examining a quadratic form that describes the current-configuration energy. Such an analysis has been performed for the equilibrium states of a vortex. At small values of the pinning parameter, the vortex state of higher energy is quasistable.

REFERENCES

1. M. A. Zelikman, Supercond. Sci. Technol. **12**, 1 (1999).
2. V. V. Bryksin, A. V. Gol'tsev, S. N. Dorogovtsev, *et al.*, Zh. Éksp. Teor. Fiz. **100**, 1281 (1991) [Sov. Phys. JETP **73**, 708 (1991)].
3. S. N. Dorogovtsev and A. N. Samukhin, Europhys. Lett. **25**, 693 (1994).
4. Ya. I. Frenkel' and T. M. Kontorova, Fiz. Zh. **1**, 137 (1939).
5. M. A. Zelikman, Zh. Tekh. Fiz. **71** (7), 44 (2001) [Tech. Phys. **46**, 831 (2001)].

Translated by A. Isaakyan

Technical Physics, Vol. 49, No. 9, 2004, pp. 1166–1175. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 63–71.

SOLID-STATE ELECTRONICS

Thermodynamics of the Growth of Carbon Nanotubes of Various Structures from Droplets of Supersaturated Melt

N. I. Alekseev

Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia

e-mail: g.dyuzhev@pop.ioffe.rssi.ru

Received September 22, 2003; in final form, January 5, 2004

Abstract—A thermodynamic model is developed to describe the growth of carbon nanotubes from a droplet of supersaturated carbon melt in a metal catalyst. The model employs molecular mechanics and semiempirical quantum chemistry methods. The results of calculations enable construction of a phase diagram showing the type of nanotube depending on the melt supersaturation with carbon and the droplet radius. © *2004 MAIK "Nauka/Interperiodica".*

Extensive application and the development of methods for the synthesis of carbon nanotubes (NTs) [1] substantially surpass understanding and theoretical modeling of the mechanisms determining the formation of a particular type of NTs. Only the growth of multiwall nanotubes (MWNTs) is described by a well-developed theory using the early model of Tibbets [2] based on the so-called vapor–liquid–droplet mechanism (VLD) [3]. According to this mechanism, NTs grow due to release of carbon from a droplet of metal melt supersaturated with carbon. The metal belongs as a rule to well-known chemical catalysts such as Fe, Fe/Ni, Y/Ni, although it was reported that NTs can be also grown using gallium as the catalyst [4]. The possibility of MWNT nucleation without catalyst was never theoretically considered.As for the single-wall nanotubes (SWNTs), there was no concept of nanotube at all when the original model [2] was proposed and, thus, they cannot be an object of investigation. The Tibbets model [2] is directly inapplicable to SWNTs. Meanwhile, the VLD mechanism is generally the only mechanism consistent with experimental data for such nanotubes. Alternative mechanisms such as catalyst scooter [5] and growth from carbon rings [6] were not apparently observed although they can most probably be realized.

Thus, despite the large body of experimental results, at present there is only a general notion of the conditions favoring the formation of a particular type of nanotubes. It is known that, in some instances, changing of the catalyst alone can result in the appearance of MWNTs instead of SWNTs. In addition, there always appear a certain amount of droplets of the catalyst encapsulated into fullerene-like cages.

A qualitative description of the formation of different types of NTs proceeding from a common standpoint was proposed by Kanzow and Ding [7]. In particular, according to a purely kinetic picture of SWNT formation given in [7], the detachment of a curved monolayer (pentagon cap) takes place if a carbon atom in the surface layer has a sufficiently high kinetic energy to overcome adhesion and surface tension of the graphite layer. The proposed scheme is quite attractive, but precludes quantitative determination of the type of nanostructure that appears.

Meanwhile, the combination of thermodynamics with the molecular mechanics (MM) calculations proposed in this study enables quantitative description of the conditions for the formation of SWNT and MWNT and for the encapsulation of catalyst droplets.

1. MODEL OF GRAPHITE ISLAND ON THE SURFACE OF DROPLET SUPERSATURATED WITH CARBON

The proposed model is restricted to the consideration of NT growth from an individual droplet. We do not consider the mechanism by which NT grows on the bridge between a small droplet with high supersaturation and a large droplet with low supersaturation [8].

The model of a nucleus was chosen so as to enable the formation of various types of NTs. The nucleus represents a portion of the graphite monolayer, appearing as an island at the interface between the metal melt and a half-space filled by an inert buffer gas (Fig. 1a). The gas acts as a thermostat maintaining equal temperature in the entire space. The central part of the island is bounded to atoms of the melt by van der Waals forces. Carbon atoms occurring at the edge of the island are connected only by two neighbor atoms of the island and are chemically bound (bold dashed lines in Fig. 1b) to metal atoms in the melt (circles in Fig. 1a). Below, these edge carbon atoms are called "radicals" to reflect the fact that their bonds to the atoms in the melt are temporary (as bonds in the melt itself). It is assumed that

1063-7842/04/4909-1166$26.00

each radical is bound to two metal atoms in the melt for most of the time, so that its valence bonds are saturated (in Fig. 1a each radical is shown bound to only one atom in the melt for the sake of clarity).

As soon as a carbon atom appears instead of one of these metal atoms, it can be extracted from the melt (atom 3 on going from Fig. 2a to Fig. 2b) and the radical (atom 1) becomes bound to three carbon atoms as in a graphite cell. If each of two neighboring radicals (atoms 1 and 2) extracts one carbon atom from the melt (atoms 3 and 4, respectively), a pentagon is formed (Fig. 2e). If one more carbon atom (atom 5 in Fig. 2c) is extracted, then a hexagon is formed (on passing from Fig. 2c to Fig. 2d). As a result, a pentagon- or hexagon-shaped ridge develops at the edge of the island.

The total binding energy of a carbon atom bound to two atoms of the iron group metals is 5–7 eV (Fig. 3a), which is approximately equal to the binding energy ε_0 of carbon in graphite. Therefore, after extraction of the carbon atom, the gain or loss of its binding energy $\Delta E_{\text{Me–C}} \ll \varepsilon_0$ is determined not quite reliably and should be varied in the course of calculations. The $\Delta E_{\text{Me–C}}$ value is determined as

$$\Delta E_{\text{Me–C}} \approx 2E_{\text{Me–C}} - E_{\text{CC}}, \qquad (1)$$

where E_{CC} and $E_{\text{Me–C}}$ are the energies of carbon–carbon bonds in the graphite lattice and carbon–metal bonds in the Me–C molecule, respectively.

After formation of the ridge, the behavior of the island depends on whether the ridge is a pentagon or hexagon. Two variants are possible if the ridge is a pentagon.

(i) The island "absorbs" a pentagon and it becomes positioned in the interior of the island. Even one pentagon inside the hexagon-packaged plane is a defect that causes the appearance of substantial stresses. When the number of pentagons N_5 becomes sufficiently large to form a pentagonal cap ($N_5 = 6$), the accumulated strain tears the central part of the graphite island from the melt boundary and the island converts into the pentagonal cap.

(ii) The appearance of a pentagon stops the island growth in the given direction. In other directions, the island continues to grow until $N_5 = 6$ in the case of flat melt boundary (or lower value in the case of finite droplet).

In order to choose between the two variants, the loss δE ($\delta E > 0$) in the energy of a flat island containing several pentagons was calculated and compared to that for the same configuration of a fragment of a fullerene surface. The island configuration was specified as follows. It was assumed that, in the absence of pentagons, hexagons are ring-clustered around the central hexagon (which is a ring with the number $n = 0$; the island with $n = 1$ corresponds to the number of hexagons $N_6 = 7$). Then, an island with one or several pentagons in the

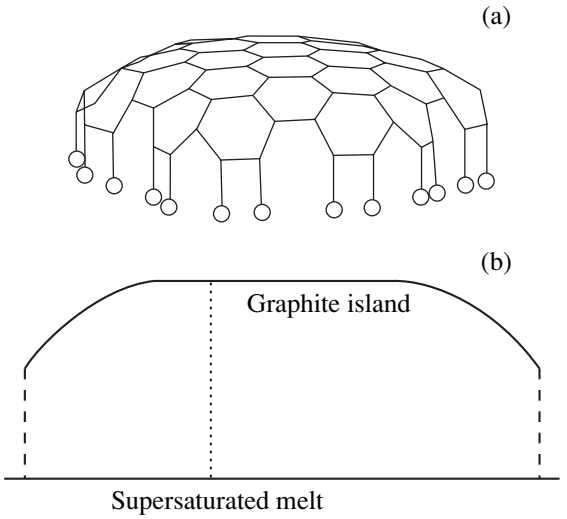


Fig. 1. Schematic diagrams showing (a) a graphite island on the melt surface and (b) its side view (circles indicate atoms of the melt; dashed lines show chemical bonds between the island atoms and atoms of the melt; dotted line shows the van der Waals bonds.

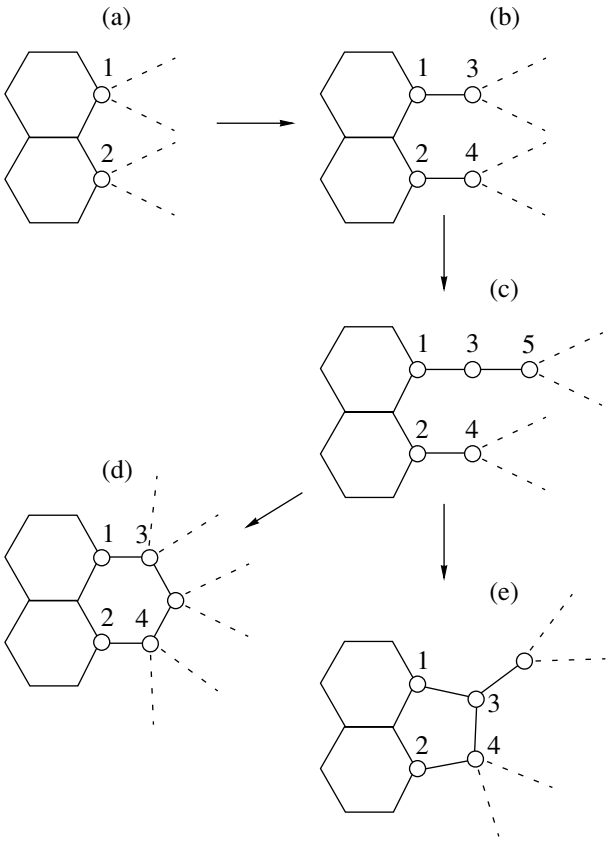


Fig. 2. The sequence of elementary reactions involved in the graphite island expansion. Dashed lines denote the bonds between island atoms and atoms of the melt. The already formed part of the island, situated to the left of pentagons and hexagons, is not shown.

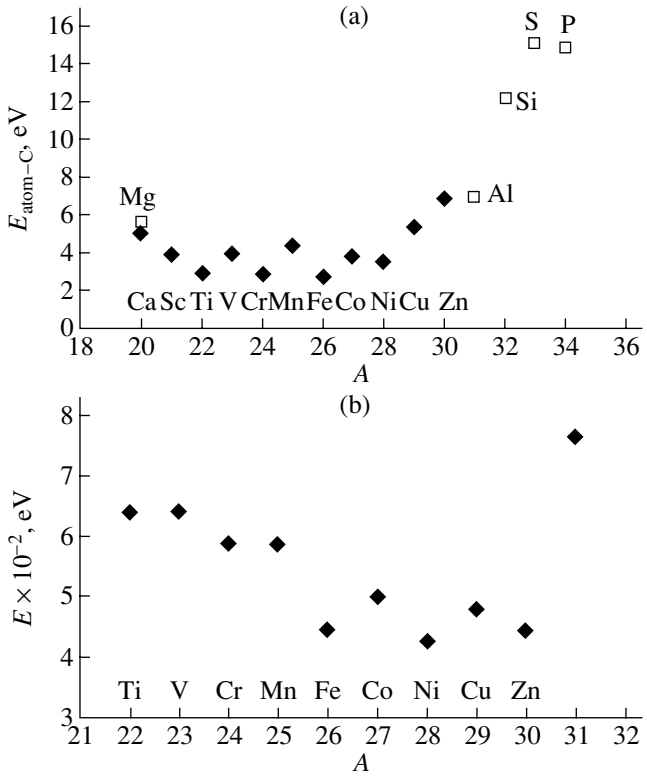


Fig. 3. Plots of (a) the binding energy of carbon atom with atoms from the first full period of the Periodic table (determined by ZINDO (◆) and PM3 (□) methods) and (b) the van der Waals interaction energy of the atom with infinite graphite plane versus atomic number of atom.

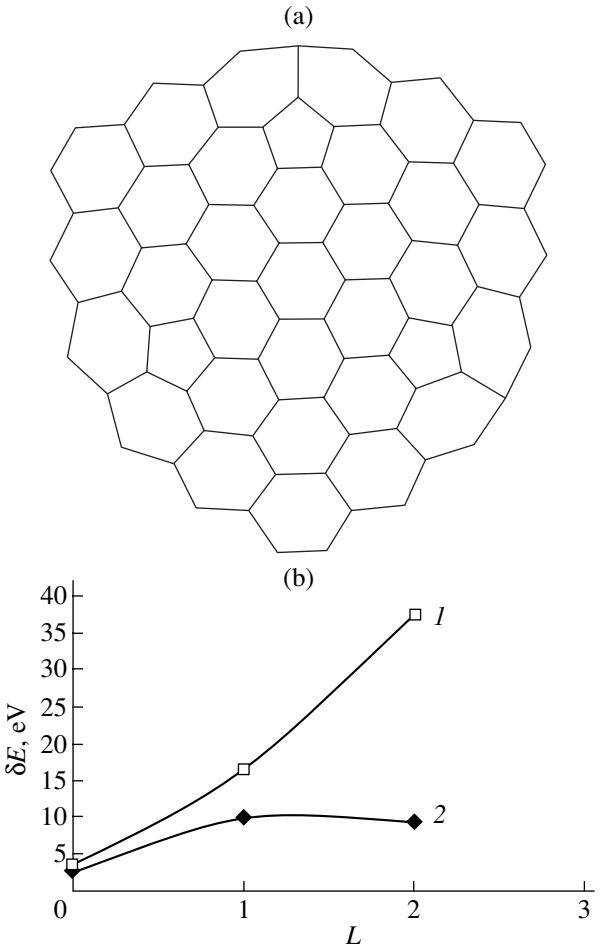


Fig. 4. (a) Schematic diagram of a flat graphite island with the number of rings of hexagons $n = 3$ and three pentagons in the second ring and (b) the dependence of the loss of binding energy δE on the position of the pentagonal defects relative to the edge of the island surface ($L = 0$ corresponds to the pentagon at the edge of the island, $L = 1$ corresponds to the pentagon in the first inner ring). The number of rings of hexagons $n = 3$; the number of pentagons $N_5 =$ (*1*) 1 and (*2*) 2.

external or in one of the internal rings (Fig. 4a) was considered.

The results of molecular mechanics (MM) calculations using the HyperChem 6.3 program package showed that δE weakly depends on the number of rings, but drastically increases with N_5 (Fig. 4b). At $N_5 = 1$ (one pentagon at the edge of the island), $\delta E \approx 2.5$ eV. However, when the pentagon is absorbed by the island, its "price" increases almost up to 10 eV. It means that the appearance of a stable pentagonal defect forces the graphite island to be "reflected" from this defect and to develop in some other direction. The possibilities of pentagon absorption by the island are not realized.

Thus, the NT emerging at the flat melt boundary should have a plateaulike cap in which all pentagons are situated at the edge. There are no "internal" pentagons. The number of internal pentagons on the curved surface of a droplet depends on the radius R_g of the droplet.

If the formed ridge is hexagon, then the bend of the rough island edge toward the melt boundary is accompanied by considerable loss of energy and formation of the overstrained structure. Therefore, on further growth of the island, the ridge should be smoothed and the edge of the island should take a form mostly closed to a circle.

Thus, the island expands via the formation of ridges and their subsequent smoothening. The binding energy oscillates and is not a convenient basis for consideration. Therefore, the barrier B associated with the formation of a regular ridge was evaluated separately (MM gives $B \sim 5$ eV), and the state of the island was fixed only in the local energy minima corresponding to round islands. As in the estimation described above, it was assumed that such an island comprises a system of rings of hexagons around a certain central element, such as an atom, bond, or hexagon.

Before the moment of detachment of the island atoms from the boundary of the melt (and the formation of a nanotube), the center of the island can move. Under

certain conditions, the detachment becomes thermodynamically favorable and the island growth ceases.

Immediately before the separation, a deficient number of pentagons $6 - N_5$ is formed at the edge of the island providing passage to the tubular part of a nanotube. Subsequent formation of one more ring of hexagons means the onset of growth of an SWNT.

Since the energetics of the island and SWNT are very similar, the formation of such NT requires rather detailed analysis. Apparently, the formation of an SWNT must occur at a small number n of rings, since the Gibbs energy gain due to release of extra carbon atoms in the tubular part of the NT is proportional to n, whereas the energy of van der Waals bonds breaking is proportional to n^2.

A different situation arises when the formation of a two-wall nanotube is considered. At a large size of the island, an almost doubled number of atoms is detached from the surface in two pentagonal caps, while the van der Waals bonds should be broken only once—for the inner wall formed inside the island. Therefore, the whole process becomes thermodynamically favorable.

2. THERMODYNAMICS OF CARBON ATOM RELEASE FROM THE MELT

2.1. Binding Energy Variation in the Case of Flat Melt Boundary

Let us begin the consideration with the case of a flat boundary of a supersaturated catalytic melt and study the Gibbs free energy variation ΔG_g upon transition of a given number g of carbon atoms from the melt into the island or what can originated from it. There are three possibilities: (i) an expanding island with the number of rings n and $N_5 = 0$ (this variant is referred to below as proliferation (PR)); (2) an SWNT including $n - 1$ rings of hexagons, pentagons in the nth ring (which provide subsequent detachment of the island atoms), and the $(n + 1)$th ring of hexagons in the tubular part of the SWNT; (3) a two-wall nanotube with a bend in the external wall in the nth ring and the tubular part of the inner wall with one ring of hexagons (the number of rings comprising the external wall can be easily calculated). Below, the two-wall nanotube is formally referred to as an MWNT.

The Gibbs free energy of the melt atoms can be expressed in terms of the energy of atoms in the melt occurring in equilibrium with graphite. Then

$$\Delta G_g = \Delta E_g - T\Delta S_g - gTw, \tag{2}$$

where ΔE_g is the loss of binding energy of the graphite island in comparison to the total energy of the same number of atoms in infinite graphite, which is determined as a positive value; ΔS_g is the loss of entropy; $w = \ln\zeta$; $\zeta = X/X_{sat}(T)$ is the supersaturation of the melt with carbon, that is, the ratio of the carbon mole fraction to the maximal steady-state value at the given temperature T.

Numerical simulation of the island at the melt boundary, especially for the curved droplet surface, is very cumbersome and rather conceals the physical essence of the problem. Therefore, the total loss of energy ΔE_g was approximately divided into several independent terms introduced "manually," and direct optimization of the system consisting of the island and atoms of the melt by MM methods was performed only for selecting the model parameters.

(i) In the case of SWNT, ΔE_g can be represented approximately as the sum

$$\Delta E_g = \Delta E_g^{(1)} + \Delta E_g^{(2)} + \Delta E_g^{(3)} + \Delta E_g^{(4)}. \tag{3}$$

The term

$$\Delta E_g^{(1)} \approx \delta_n/2 \tag{4}$$

accounts for the loss in the binding energy due to distortion of the valence angles of carbon bonds in the island. Since the bonds between radicals and atoms of the melt form a nearly right angle with the melt boundary, the island can be identified with the upper half of the fullerene in which all pentagons are grouped in two adjacent rings along the equator, while the rest of the surface is almost flat. The dependence of the loss of binding energy in such fullerene, δ_n, on the number of rings n in each of its halves is well described for $n \geq 1$ by the function

$$\delta_n = 2\lambda n + \alpha \tag{5}$$

and has clear physical meaning. For a large number of atoms, the loss of energy consists of the component α related to pentagonal distortions and the term describing the type of edge distortion proportional to the length of the island edge. The parameters λ and α obtained as a result of geometry optimization using MM method are approximately $2\lambda = 5.0$ eV, $\alpha = 10.3$ eV. The term

$$\Delta E_g^{(2)} \approx \rho_n \Delta E_{\text{Me–C}} \tag{6}$$

takes into account that the radicals at the island edge are bound to two metal atoms, rather than to carbon atoms, ρ_n is the number of radicals. The topological factors g and ρ_n are given by the formulas

$$\begin{aligned} g &= 6(n^2 + \beta_{SW} n + \gamma_{SW}), \\ \rho_n &= 6(n + \varphi_{SW}), \end{aligned} \tag{7}$$

where the parameters β_{SW}, γ_{SW}, φ_{SW} depend on the type of the island center. For example, if the center is hexagon and the number of rings of hexagons is $n = 0$, then $\beta_{SW} = 4$, $\gamma_{SW} = 0$, $\varphi_{SW} = 0$.

The term $\Delta E_g^{(3)}$ accounts for the distortions of the valence angle of the carbon bonds in the tubular part of the NT as compared to the graphite. Molecular mechanics calculations showed that $\Delta E_g^{(3)}$ weakly depends on the NT diameter and amounts (per ring of

hexagons) approximately to

$$\Delta E_g^{(3)} = \alpha_{VT} \approx 0.9 \text{ eV}. \quad (8)$$

The term $\Delta E_g^{(4)}$ stems from the rupture of van der Waals bonds between island atoms and metal atoms in the melt and can be expressed through the van der Waals interaction $A > 0$ of one metal atom of the given kind with an infinite graphite plane (Fig. 3)

$$\Delta E_g^{(4)} = A g_{n-1}, \quad (9)$$

where g_{n-1} is the number of carbon atoms in the internal part of the island.

Finally

$$\Delta G_g^{(SWNT)} = \frac{\alpha}{2} + \lambda n + 6n\Delta E_{Me-C} + \alpha_{NT} + 6An^2 - 6n(n+4)w - T\Delta S_g. \quad (10)$$

(ii) In the case of proliferation (PR), $\Delta E^{(1)}$ is similar to (2), but with replacement of $\alpha/2$ by a different constant $\alpha_1 \approx 4.2$ eV (the calculation of α_1 is also possible using some modification of the HyperChem 6.3 package). The term $\Delta E^{(2)}$ is similar to (5), but the values β_{SW}, γ_{SW}, and φ_{SW} in (7) should be replaced by $\beta_{PR} = 2$, $\gamma_{PR} = 1$, and $\varphi_{PR} = 1$, respectively.

Finally,

$$\Delta G_g^{(PR)} = \alpha_1 + \lambda n + 6(n+1)\Delta E_{Me-C} - 6(n+1)^2 w - T\Delta S_g. \quad (11)$$

(iii) In the case of the two-wall NT (i.e., MWNT), $\Delta G_g^{(MWNT)}$ is calculated similar to (10) and (11) and is determined by the relationship

$$\Delta G_g^{(MWNT)} = \alpha_1 + \lambda(n+n_2) + 6(n+n_2)\Delta E_{Me-C} + 2\alpha_{NT} - 6An_2^2 - (6n_2(n_2+4) + 6n(n+4) + 2\rho_n\varepsilon)w - T\Delta S_g, \quad (12)$$

where $\varepsilon = h/a\Psi$ is an additional number of rings in the pentagonal cap and the tubular part of the outer wall of the NT as compared to the inner one; $a\Psi$ is the width of one ring of hexagons; $a \approx 1.5$ Å is the length of the C–C bond; $0.5 \le \Psi \le 1.8$ is a variable parameter; $n_2 = n - \varepsilon$ is the number of rings in the inner wall; $\rho_n = 6n$ is the number of radicals in the bend ring of the outer wall.

2.2. Entropy Loss Evaluation

The entropy of the graphite island as an ensemble of harmonic oscillators is determined by summing over all oscillators i

$$S_g = -\sum_i \ln(1 - \exp(-h\nu_i/T)). \quad (13)$$

At a high temperature, $S_g \approx -\sum_i \ln(h\nu_i/T)$. Estimations showed that, for the average oscillator frequency $\langle\nu\rangle = 5 \times 10^{-12}$ s^{-1} and $T = 0.2$ eV, we have the ratio $h\langle\nu\rangle/T \sim 0.1$. As the binding energy of carbon atoms in graphite is $\varepsilon_0 = 6$ eV, the energy term in G_g is approximately $9g$ per atom, whereas the entropy term is $\sim 1g$. Thus, the entropy term is almost one order of magnitude lower and can be omitted in rough estimations.

In the approximation of average oscillators frequency, the loss of entropy ΔS_g is

$$\Delta S_g \approx -(3g-6)\ln\langle\nu\rangle/\nu_0, \quad (14)$$

where ν_0 is the average vibrational frequency in graphite.

The calculation of the vibrational spectrum of the island requires enormous computational time; for this reason, the average frequency was estimated as follows. All atoms were divided into three groups: inner atoms, atoms at the bend of the island, and radicals. In the island of PR type consisting of n rings with the total number of atoms $g = 6(n+1)^2$, the corresponding numbers of atoms are $6n^2$, $6(2n+1) \approx 18n$, and $6(n+1)$, respectively. The number of degrees of freedom per atom is $3(1 - 2/g)$.

Then the ratio $\langle\nu\rangle/\nu_0$ in (14) can be estimated from the relationship

$$\left(\frac{\langle\nu\rangle}{\nu_0}\right)^{3g-6} = \left[\left(\frac{\nu_0'}{\nu_0}\right)^{18n^2}\left(\frac{\nu_{bend}}{\nu_0}\right)^{36n}\left(\frac{\nu_{Me-C}}{\nu_0}\right)^{1-2/g}\right], \quad (15)$$

where ν_0', ν_{bend} are the average frequencies of atomic vibrations in the flat part of the island and the bend region, respectively, and ν_{Me-C} is the vibration frequency of the bond between radical and metal atom in the melt.

The second and thirds multipliers in (15) are the most important. The frequency ν_{bend} estimated using AM1 method was 25–30% higher than the frequency ν_0: $\nu_{bend}/\nu_0 = 1.3$. The frequency ν_{Me-C} was estimated from the vibration frequency of the FeC molecule calculated using ZINDO method: $\nu_{Me-C}/\nu_0 = 1/1.4$.

2.3. Variation of the Binding Energy at the Surface of the Melt Droplet

In the case of a melt droplet of finite radius, the inner part of the island involves pentagons, which can be distributed in different ways. In this study, we assume that these pentagons are distributed uniformly and the island can be considered as a single object. Two approaches can be applied for the description of such an island.

(i) It is possible to consider the inner part of the island as a fragment of spherical fullerene with elements typical of fullerenes such as pentagons, hexagons, etc., whose numbers are proportional to the rela-

tive area of the spherical segment of the $(n-1)$th ring. The number of pentagons in the bend ring with the number n is six in the case of PR and supplement to six in the SWNT and MWNT cases. With this approach, the energy loss components related to the bend ring and inner part of the island are clearly separated.

(ii) It is possible not to divide the island into the bend ring and inner part, but determine ΔG_g through interpolation between two extreme cases: small island ($n \longrightarrow 0$) and very large island, which covers half of the droplet surface of radius R_g and is bent to its equator (Fig. 5). The last approach was used due to its simplicity. For the small island, all equations derived above are valid. The island leaning on the droplet equator (with the number of rings $n = n_0$) is special: the inner $n_0 - 1$ rings of the island include already almost six pentagons, whereas t here are almost no pentagons in the region of the bend. Therefore, the cases of SWNT and PR at $n = n_0$ are distinguished only by the presence of one ring of hexagons in the tubular part of NT for SWNT.

The loss of energy for SWNT in the bend ring is the same as for a hypothetical island *ABCD* of radius $R_F = R_g + h$ (where $h \approx 3.34$ Å is the distance from the island to the melt boundary that is assumed to be equal to the distance between the basal planes in graphite), which leans on the *AB* plane (Fig. 5). Such an island conforms to the flat melt boundary and its energetic scan be considered as known.

The loss of binding energy per ring of hexagons in the tubular part of NT, α, as well as in the NT leaning on the flat melt boundary, was assumed to be α_{NT}. Thus, the values $\Delta E^{(1)}$, $\Delta E^{(2)}$ at $n = n_0$ for SWNT and PA are

$$\Delta E^{(1)} = \lambda n_U + \alpha'/2 + \alpha_1, \quad (16)$$

$$\Delta E^{(2)} = \rho_0 \Delta E_{\text{Me–C}}, \quad (17)$$

where $n_U = R_F/a\Psi$ is the number of rings in the island *ABCD* (Fig. 5); $\alpha' = 9.6$ eV is the loss of binding energy of a round fullerene of radius R_F as compared to graphite (which is almost independent of the size of the fullerene);

$$\rho_0 = 2\pi R_F/a\Psi \quad (18)$$

is the number of radicals formed upon cutting fullerene of radius R_F across equator, which is approximately equal to the number of hexagons in the section.

In the case of PR with $n = n_0$, the number of atoms is approximately

$$g_0 = N_F^{(\text{round})}/2 + \rho_0, \quad (19)$$

where $N_F^{(\text{round})}$ is the number of atoms in a round fullerene of radius R_F.

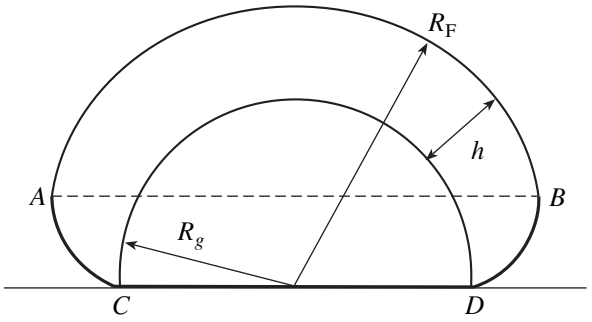


Fig. 5. Graphite island leaning on the equator of a droplet (d) of melt. The dashed line *AB* denotes the position of the melt surface on which the imaginary flat island *ABCD* is leaning.

The values $\Delta E^{(3)}$ and $\Delta E^{(4)}$ for SWNT are the same as in the case of flat melt boundary; the g_{n-1} value was assumed to be equal to $g_{n-1} = g_0 - 2\rho_0$.

The energy loss components for MWNT with $n \longrightarrow n_0$ were evaluated in the same manner. The external wall was identical to SWNT, but the extra number of rings in the tubular part released on formation of one tubular ring of the inner ring was taken into account. Formulas for the inner wall with the number of rings $n_0 - \varepsilon$ were obtained by interpolation between the equations for SWNT with $n \longrightarrow 0$ and $n = n_0$

$$f = f(n \longrightarrow 0)(1 - J(n/n_0)) + \vartheta(n/n_0)(f(n_0) - f(n \longrightarrow 0)).$$

The form of the function $\vartheta(n/n_0)$ was varied.

For islands and nanotubes of intermediate size, $0 \le n \le n_0$, the components of the loss of binding energy and the topological factors were also determined by interpolation between the formulas for the flat island $n \longrightarrow 0$ and spherical droplet $n = n_0$.

3. RESULTS AND DISCUSSION

3.1. Flat Melt Boundary

Figure 6 shows the Gibbs free energy as a function of the number of rings n in the case of formation of PR graphite islands and the subsequent growth of nanotubes. The variable parameter $\Delta E_{\text{Me–C}}$ was taken equal to $\Delta E_{\text{Me–C}} = -0.2$.

Among the three curves presented in Figs. 6a and 6b, only the PR curve can describe the development of the island. The actual phase trajectory of the island before the moment of detachment starts in the PR curve and moves down gradually as pentagons are formed at the edge of the island. However, it is impossible to say how many pentagons are already formed for a given n and, hence, the form of the real trajectory is not determined unambiguously. Therefore, it was assumed that before detachment of the island atoms from the melt

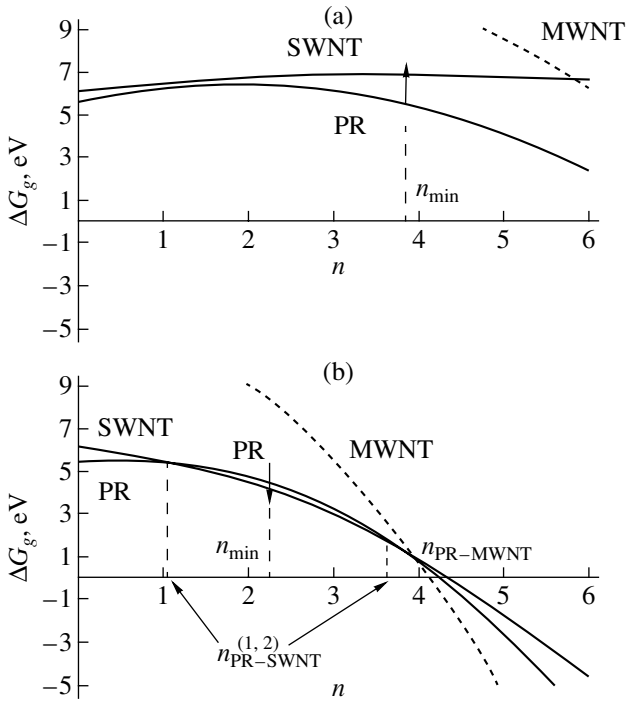


Fig. 6. Plots of the Gibbs energy variation ΔG_g versus the number of rings in the PR, SWNT and MWNT cases for the carbon supersaturation $\zeta = 1.3$ (a), 1.6 (b).

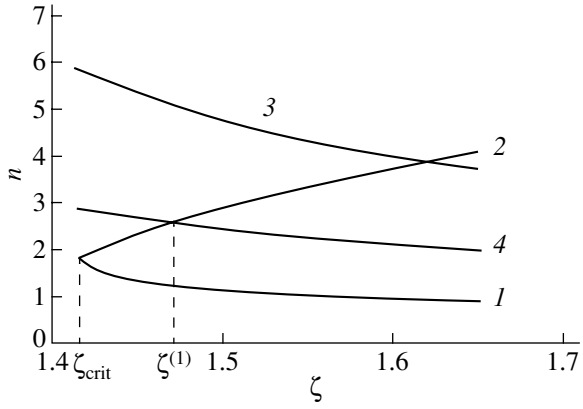


Fig. 7. Dependence of the points of intersection $n_{PR-SW}^{(1)}$, $n_{PR-SW}^{(2)}$ (curves *1* and *2*, respectively), n_{PR-MW} (*3*) and minimal size of the nanotubes n_{min} (*4*) on the supersaturation ζ. The values of parameters $\alpha_1 - \alpha/2 = 1.8$ eV, $\Delta E_{Me-C} = -0.2$ eV.

boundary, this island develops along to the PR trajectory. The island cannot move along the SWNT and MWNT curves. They are necessary only for determining the type of the growing object.

Obviously, for large *n*, the MWNT curve is below the PR curve because at $n \longrightarrow \infty$ the multiplier at n^2 for the two-wall NT is approximately twice as large as that for SWNT. Therefore, the point n_{PR-MW} of the intersection of PR and MWNT curves always exists. This means that, for an unlimited time of experiment, an MWNT always appears, unless an SWNT is formed at a smaller *n*.

The SWNT emerges if the corresponding curve is below the PR curve in a certain range of *n*, that is, of the point of intersection n_{PR-SW} exists. Moreover, the possibility of jumping from the PR curve to the SWNT curve is determined by two additional factors.

(i) As the island (PR) expands by one ring, a barrier *B* for the appearance of a hexagonal ridge should be overcome (Fig. 2). This makes possible the jump from PR to SWNT even when the SWNT curve is lying higher, but the barrier that needs to be jumped to reach it is smaller than *B* (Fig. 6a, arrow indicating upward). However, the calculation shows that such a situation is realized only in a narrow range of supersaturations ζ and requires detailed additional consideration, which is beyond the accuracy of the approach developed here.

(ii) The development of an SWNT emerging after detachment of the island from the melt boundary must be thermodynamically favorable. Therefore, the gain achieved when additional atoms are released from each ring in the tubular part of the NT should exceed the loss in the binding energy resulting from their getting into the tubular part of the NT rather than onto the graphite plane. Thus, the condition $\rho_n T \ln\zeta - \alpha_{NT} > 0$ must be fulfilled, which establishes the minimal nanotube radius (with allowance for (8)) for each ζ:

$$n_{min} = \alpha_{NT}/6T\ln\zeta. \qquad (20)$$

Analysis of the family of curves $\Delta G_g^{(PR)}$, $\Delta G_g^{(SWNT)}$, and $\Delta G_g^{(MWNT)}$ shows that, for small ζ, the PR curve lies below the SWNT curve (Fig. 6a). This means that the SWNT curve is forbidden and an MWNT emerges when the island expands to a size corresponding to the point of intersection n_{PR-MW} of the PA and MWNT curves.

As the supersaturation increases to a certain value ζ_{crit}, the SWNT curve touches the less sloped PR curve (Fig. 6b), so that at $\zeta > \zeta_{crit}$ there are two points of intersection, $n_{PR-SW}^{(1)}$ and $n_{PR-SW}^{(2)} > n_{PR-SW}^{(1)}$. In the range $n_{PR-SW}^{(1)} \le n \le n_{PR-SW}^{(2)}$, the jump from PR to SWNT and the formation of SWNT is possible, provided the condition $n > n_{min}$ is fulfilled (Fig. 6b, bold arrow pointing downward). The dependences of the characteristic points of intersection on ζ are shown in Fig. 7. As can be seen for the selected parameters of calculation, the transition through the critical supersaturation is not a sufficient condition for the SWNT formation: if ζ exceeds ζ_{crit}, but is lower than $\zeta^{(1)}$, then $n_{min} \ge n_{PR-SW}^{(2)}$ and the jump PR $\rightarrow$ SWNT is still impossible. At $\zeta > \zeta^{(1)} n_{PR-SW}^{(1)} < n_{min} < n_{PR-SW}^{(2)}$ and an SWNT of size n_{min} is formed.

Since the parameter $\Delta E_{\text{Me–C}}$ is variable, the dependence of the results on this value should be checked. This is especially important for the points of intersection of the PR and SWNT curves.

If we neglect for simplicity the entropy variation in (10) and (11), then the expression for PR retains one term proportional to n^2: $\Delta G_g^{(PR)}(n \longrightarrow \infty) \sim -6wn^2$. In the case of SWNT, there appears an additional term related to the rupture of the van der Waals bonds : $\Delta G_g^{(SWNT)}(n \longrightarrow \infty) \sim 6(A - w)n^2$.

The condition of the intersection of SWNT and PA curves has the form

$$6An^2 - 12nw - (\alpha_1 - \alpha/2 + 6\Delta E_{\text{Me–C}} - \alpha_{\text{NT}}) = 0. \quad (21)$$

A single solution of the quadratic equation (21) exists when

$$6w^2 - 6Aw + A(\alpha_1 - \alpha/2 + 6\Delta E_{\text{Me–C}} - \alpha_{\text{NT}}) = 0. \quad (22)$$

The solution of (22) is

$$w = \frac{3A \pm \sqrt{9A^2 - 6A(\alpha_1 - \alpha/2 + 6\Delta E_{\text{Me–C}} - \alpha_{\text{NT}})}}{6} \quad (23)$$

under the condition that its discriminant is non-negative. When

$$3A > 2(\alpha_1 - \alpha/2 + 6\Delta E_{\text{Me–C}} - \alpha_{\text{NT}}) \quad (24)$$

there is no solution for w and, regardless of the supersaturation, there will apear either only SWNT or only MWNT.

Since the adhesion A is two orders of magnitude lower than each of three terms in the right-hand side of (24), the inequality (24) means in fact that

$$\alpha_1 - \alpha/2 + 6\Delta E_{\text{Me–C}} - \alpha_{\text{NT}} < 0. \quad (25)$$

If $\alpha_1 - \alpha/2 + 6\Delta E_{\text{Me–C}} - \alpha_{\text{NT}} > 0$, then there exists a certain w for which the formation of an SWNT changes to the growth of MWNT. For the chosen parameters, this inequality is fulfilled and this fact determines the above results. For $\alpha_1 - \alpha/2 = 1.8$ eV, as assumed in calculations, the results are qualitatively valid up to $\Delta E_{\text{Me–C}} \le -0.15$ eV.

The dependence of the points of intersection $n_{\text{PR–SW}}^{(1)}$, $n_{\text{PR–SW}}^{(2)}$, $n_{\min}$, $n_{\text{PR–MW}}$ on supersaturation in the opposite case (i.e., when inequality (25) is fulfilled) is shown in Fig. 8. The calculation parameters were chosen as $\alpha_1 - \alpha/2 = 1.8$ eV, $\Delta E_{\text{Me–C}} \le +0.1$ eV. In this case, $n_{\text{PR–MW}}^{(1)} < 0$ and is not shown here, whereas $n_{\min} < n_{\text{PR–SW}}^{(2)} < n_{\text{PR–MW}}$ over a wide range of ζ values. Therefore, an SWNT of size $n_{\min}$ should be formed for any ζ.

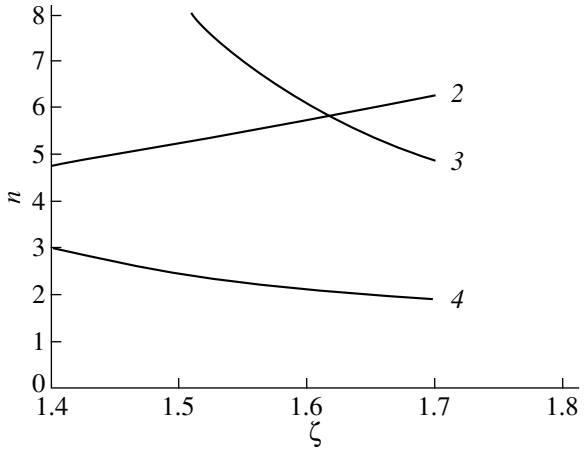


Fig. 8. The same dependencies as in Fig. 7 for the values of parameters $\alpha_1 - \alpha/2 = 1.8$ eV, $\Delta E_{\text{Me–C}} = +0.1$ eV.

The maximum of the free energy barrier for NT nucleation depends on the specific features of the intersection of the PA and SWNT curves, but it is close to the maximum of the SWNT curve, that is,

$$\Delta G_{\max} \cong \frac{\alpha}{2} + \alpha_{\text{NT}} + \frac{(\lambda + 6\Delta E_{\text{Me–C}})^2}{24(w - A)}. \quad (26)$$

Obviously, under otherwise equal conditions, $\Delta G_{\max}$ decreases with decreasing A. On the other hand, $\Delta E_{\text{Me–C}}$ should not be "too negative." The optimal value of $\Delta E_{\text{Me–C}}$ is

$$\Delta E_{\text{Me–C}} \sim -\lambda/6 = -0.3 \text{ eV}. \quad (27)$$

The dependence of the binding energy of metal atoms to carbon and similar dependences of the energy of the van der Waals interaction of the metal atoms with the graphite plane (Fig. 3b) show that the transition metal atoms are optimal from the point of view of detachment of the island atoms from the melt and NT formation. Their cohesion with the graphite plane is minimal and the binding energy is most close to the value that provides fulfillment of condition (27).

If the melt consists of two metals, M_1 (with a mole fraction of X_1) and M_2, the barrier $\Delta G_{\max}$ takes the form

$$\Delta G_{\max} \cong \frac{\alpha}{2} + \alpha_{\text{HT}} + \frac{(\lambda + 6(X_1\Delta E_{\text{Me–C}}^{(1)} + (1 - X_1)\Delta E_{\text{Me–C}}^{(2)}))^2}{24(w - A_1X_1 - A_2(1 - X_1))}. \quad (28)$$

This expression has a minimum at the mole fraction

$$X_1 = \frac{12w(\Delta E_{\text{Me–C}}^{(1)} - \Delta E_{\text{Me–C}}^{2}) + \lambda(A_1 - A_2) + 6(A_2(\Delta E_{\text{Me–C}}^{(2)} - 2\Delta E_{\text{Me–C}}^{(1)}) + A_1\Delta E_{\text{Me–C}}^{(2)})}{6(A_1 - A_2)(\Delta E_{\text{Me–C}}^{(1)} - \Delta E_{\text{Me–C}}^{2})}, \quad (29)$$

which determines the optimal ratio between the catalysts in the melt.

3.2. Nucleation of NT at the Surface of Finite Droplet

On going to the case of a curved melt surface, the shape of the plots presented in Fig. 6 somewhat changes quantitatively and a qualitative feature appears that is indicative of an additional possibility.

Obviously, the SWNT and MWNT curves are meaningless on the right side from the point n_0 corresponding to the equator of the droplet. Therefore, if at $n = n_0$ an SWNT does not appear and the MWNT curve is above the PR curve, then the graphite island crosses the equator and the droplet is encapsulated into a fullerene-like cage. Below, the same notation PR is used for the encapsulated droplets of the catalysts as for the process of island proliferation.

Simple qualitative analysis of the behavior of MWNT and PR curves for a large size of droplet, analogous to that presented above for the SWNT and PR curves, shows that an MWNT cannot be formed at a level of supersaturation of $\ln\zeta < A/T$. Higher supersaturation is required to form MWNTs on the droplets of smaller size.

For very small droplets, the growth of both SWNTs and MWNTs is impossible and nothing except droplet encapsulation can take place. Thus, the type of the nanotube can be shown in the diagram of R_g plotted against supersaturation ζ. Figure 9 is plotted for $T = 0.15$ eV and $\Delta E_{\text{Me–C}} = -0.2$ eV, so that both SWNTs and MWNTs can appear. As the ordinate axis, it is also possible to use the product $T\ln\zeta$. However, the diagram in fact depends on ζ and T separately since the temperature appears not only in the product $T\ln\zeta$, but in the entropy term as well. Nevertheless, there is not much sense in introducing the third axis taking into account the qualitative character of the results.

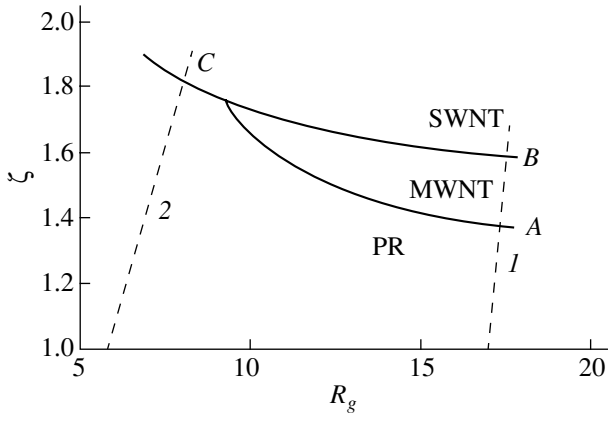


Fig. 9. Phase diagram of the single-wall nanotubes (SWNT)—encapsulation (PR)—multi-wall nanotubes (MWNT) system.

4. DEPENDENCE OF THE NANOTUBE TYPE ON EXPERIMENTAL CONDITIONS

Depending on the method of NT fabrication, the droplet supersaturation with carbon ζ varies with time t according to a certain law. Let the droplet become saturated at the instant $t = 0$ ($\zeta = 1$). Then the system behavior with increasing ζ depends on the droplet size. At a large radius of the droplet (Fig. 9, dashed line *1*), the number of droplets encapsulated as ζ changes from the point O_1 to A is negligible. An increase in the concentration n_{MWNT} of MWNTs with ζ changing from A to B is described by the equation

$$dn_{\text{MWNT}}/dt = (N_a T/\hbar)\exp(-\Delta G_g^{(\text{MWNT})}/T),$$

where N_d is the concentration of droplets.

The relative amount of MWNTs is

$$\frac{\delta N_{\text{MWNT}}}{N_d} = \frac{T}{\hbar}\int_{t(A)}^{t(B)} dt\exp(-\Delta G_g^{(\text{MWNT})}(\zeta(t))/T)$$

$$= \frac{T}{\hbar}\int_{\zeta(A)}^{\zeta(B)} d\zeta(d\zeta/dt)^{-1}\exp(-\Delta G_g^{(\text{MWNT})}(\zeta)/T).$$

If $\delta N_{\text{MWNT}}/N_d \sim 1$, then almost all droplets give rise to MWNTs. If $\delta N_{\text{MWNT}}/N_d \ll 1$, then mainly SWNTs are generated. For smaller droplets (curve *2*), either substantial number of encapsulated droplets with a small number of SWNTs, or only SWNT can appear.

5. THE POSSIBILITY OF COMPARISON WITH EXPERIMENT

Since the values $\Delta E_{\text{Me–C}}$, λ, and α, α_1 entering into Eqs. (4)–(11) are not known with sufficient accuracy, the barrier $\Delta G_{\max}$ and the growth rate of NT indeed cannot be calculated. Therefore, comparison of calculations using Eq. (28) with the experimental data available is presently impossible.

The most serious verification of the developed model is related to the calculation of the optimal mixture of catalysts using (29) and comparison of the results to the behavior of mixtures selected experimentally. Such verification will be performed in a separate study.

6. CONCLUSIONS

Combination of the methods of thermodynamics and molecular mechanics enables prediction of the type of a nanostructure that appears as carbon is released from a droplet of a metal catalyst supersaturated with carbon. The type of the structure for a given catalyst depends on the droplet radius R_g, temperature T and supersaturation ζ of the melt with carbon.

Nanostructures appearing when carbon is released from the melt form the following zones in the $R_g - T\ln\zeta$ diagram: single-wall NTs (SWNT)—multiwall NTs (MWNT)—encapsulation (PR).

Depending on the parameters of the catalyst, the diagram can include either all three zones, or only MWNT and PR zones. For the known catalysts (transition metals) all three zones are realized, and their qualitative arrangement corresponds to that depicted in Fig. 9, in particular, when the droplet radius $R_g \longrightarrow \infty$ and the supersaturation is small, then MWNTs are formed, whereas at high supersaturations, SWNTs appear. For the droplets of finite radius, the zone of droplet encapsulation into the fullerene-like cage emerges.

REFERENCES

1. *Perspective of Fullerene Nanotechnology*, Ed. by E. Osawa (Kluwer, Dordrecht, 2002).
2. G. O. Tibbets, J. Cryst. Growth **66**, 632 (1984).
3. E. I. Givargizov, *Vapor Growth of Filamentary and Scaly Crystals* (Nauka, Moscow, 1977) [in Russian].
4. Z. W. Pan, S. Dai, D. Beach, *et al.*, Appl. Phys. Lett. **82**, 1947 (2003).
5. Y. H. Lee, S. G. Kim, and D. Tomanek, Phys. Rev. Lett. **78**, 2393 (1997).
6. C. H. Kiang and W. Goddard III, Phys. Rev. Lett. **76**, 2515 (1996).
7. H. Kanzow and A. Ding, Phys. Rev. B **60**, 11180 (1999).
8. D. C. Li, L. Dai, and S. Huang, Chem. Phys. Lett. **316**, 349 (2000).

Translated by M. Lebedev

Technical Physics, Vol. 49, No. 9, 2004, pp. 1176–1180. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 72–76.

SOLID-STATE ELECTRONICS

Relaxer Ferroelectrics as Promising Materials for IR Detectors

S. E. Aleksandrov, G. A. Gavrilov, A. A. Kapralov, E. P. Smirnova, G. Yu. Sotnikova, and A. V. Sotnikov

Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia

e-mail: gga@holo.ioffe.rssi.ru

Received February 17, 2004

Abstract—The dielectric and pyroelectric properties of a typical relaxer ferroelectric, $0.9PbMg_{1/3}Nb_{2/3}O_3$–$0.1PbTiO_3$ (PMN–PT), are studied experimentally. Based on the results obtained, the pyroelectric constant and figure of merit of the material when used in IR detectors are calculated. These parameters are presented as a function of temperature and external electric field. The current and voltage sensitivities and the detectivity of PMN–PT-based IR detectors are evaluated. They are compared with the same properties of pyroelectric detectors and dielectric bolometers that use traditional pyroelectric materials as the active element and also of other uncooled photodetectors. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The use of IR radiation in a variety of devices (gas and flame detectors and analyzers, fine-vision devices, motion detectors, pyrometers, etc.) is closely related to the advances in IR radiation detection. Such factors as cost and possibility of fabricating large-area detectors are no less important than the radiometric characteristics of the devices and often dictate the choice of the detector. Among the most important requirements for these devices is reliable operation without forced cooling.

Photodetectors are usually compared in terms of their detectivity D^*, which is defined as the reciprocal of the least detectable signal power P_N: $D^* = \sqrt{A}/P_N$, where A is the photodetector area [1]. Since P_N depends on the sensitivity and noise of the photodetector (the latter parameter being proportional to the square root of the photodetector area in most cases), $D^* = \sqrt{A}/P_N$ is independent of the photodetector area and may be used for comparing the performance of photodetectors of different dimensions.

Conventional photodetectors intended for the mid-IR range (up to 10.6 μm), such as semiconductor (InSb or $Hg_{1-x}Cd_xTe$) photodiodes, offer a detectivity as high as $D^* = 10^{10}$ cm/(Hz$^{1/2}$ W) only at cryogenic temperatures. PbSe photoresistors, which are in common use, have D^* on the order of (10^8–10^9) cm/(Hz$^{1/2}$ W) with a sharp dip near $\lambda = 4.5$ μm, which shifts toward shorter wavelengths with increasing temperature [2]. Photodiodes based on III–V compounds (the devices have been recently developed at the Ioffe Physicotechnical Institute) operate at room temperature and, according to our estimates, provide D^* in the range (10^8–5×10^9) cm/(Hz$^{1/2}$ W). However, they demonstrate a high sensitivity only in narrow spectral intervals $\Delta\lambda$ (0.5 μm) between 2.9 and 5 μm [3].

By the principle of operation, uncooled heat detectors based on pyroelectric materials have the same sensitivity throughout the electromagnetic spectrum and find wide application mostly as infrared detectors. The energy absorbed by the pyroelectric element changes its temperature, which, in turn, changes its polarization and, accordingly, induces surface charge. The pyroelectric effect is observed in all media that possess a preferential polarization direction. One of the basic factors limiting the application of pyroelectric materials in radiation detectors is that the pyroelectric current, by its physical nature, is the response of the system only to temperature changes; therefore, for the radiation to be detected, it must be modulated.

The materials most widely used in this field are triglycinesulfate (TGS) and TGS-based compounds, which offer the highest sensitivity: their typical detectivity, 5×10^8 cm/(Hz$^{1/2}$ W) [4], is comparable to that of uncooled photodiodes. However, TGS crystals are rather brittle and hygroscopic, as well as have a very low thermal conductivity.

The pyroelectric effect can also be observed in nonpolarized materials on application of a dc electric field. In this case, we deal with the so-called field-induced pyroelectric effect, and detectors made of these materials are called dielectric bolometers. They have a somewhat lower detectivity ($D^* = (2–5) \times 10^8$ cm/(Hz$^{1/2}$ W) [5–12]) than standard pyroelectric detectors but are stable to high levels of illumination and, what is most important, allow one to replace radiation modulation by applied field modulation. This stimulates interest in

1063-7842/04/4909-1176$26.00 © 2004 MAIK "Nauka/Interperiodica"

seeking materials for pyroelectric IR detectors that operate as dielectric bolometers.

In recent years, attention has been focused on ferroelectrics with a diffuse phase transition (relaxer ferroelectrics). The distinctive features of relaxers are the ferroelectric transition occurring in a wide temperature range and the possibility of varying this temperature range within wide limits by modifying the composition of the material. As a representative of relaxer ferroelectrics, one should first mention $PbMg_{1/3}Nb_{2/3}O_3$ (PMN) and related solid solutions. Note that it is basically impossible to reach the polarized state in the range of the diffuse phase transition without a dc electric bias, because the temperature will break an arising polarization direction. Thus, the field-induced pyroelectric effect alone can be observed in these materials. This effect is still not fully understood. Moreover, the pyroelectric constant is hard to evaluate adequately in experiments. The main experimental difficulty is to pick up the weak induced pyroelectric current from thermally stimulated and leakage currents, which appear in the samples exposed to a dc biasing field. Finally, it is particularly important to study the material under conditions as close to those under which IR detectors operate as possible. Among many methods of measuring the pyroelectric response, the modulation method (or the Chainovis method) and its modifications [4–6, 13, 14] meet the requirements to the largest extent.

The goal of this work is to study the pyroelectric effect in PMN–PT, a typical relaxer ferroelectric, and predict the performance of the IR detectors that may be built around it. To experimentally study relaxer ferroelectrics, we developed a computerized instrument that is capable of generating the pyroelectric current in various regimes, extracting weak currents when a high dc voltage is applied, controlling the parameters of the sounding radiation, and varying the sample temperature from 278 to 350 K [15]. In the experiments, we used the solid solution of Mg-substituted lead niobate and lead titanate (PMN–PT). Based on the measured insulating and pyroelectric properties of these materials, their figures of merit were calculated and compared with those of other pyroelectric materials and IR detectors.

EXPERIMENTAL RESULTS

The pyroelectric constant p of a material subjected to a dc biasing field E is defined as [4, 6, 8]

$$p \cong \frac{\partial P_s}{\partial T} + \varepsilon_0 \int_0^E \frac{\partial \varepsilon(E, T)}{\partial T} dE, \qquad (1)$$

where P_s is the spontaneous polarization, ε is the permittivity of the material, E is the electric field intensity, and T is the temperature.

The first term on the right-hand side of the equation is the pyroelectric constant of polar materials that is due to the spontaneous polarization P_s, and the second term describes the so-called induced pyroelectric effect, which is observed in nonpolarized ferroelectrics in the presence of an electric bias. If the temperature of a pyroelectric element exposed to radiation of power $W(t)$ that is modulated by the law $W(t) = W_0\{1 + \exp(i\omega t)\}$ changes by ΔT, an expression for the pyroelectric current density will contain a component with a frequency ω:

$$J(\omega) = p\frac{dT}{dt} = pi\omega\Delta T\exp(i\omega t). \qquad (2)$$

Thus, measuring the pyroelectric current in the sample and knowing ΔT, one can calculate the pyroelectric constant of the material.

It is customary to compare pyroelectric materials in terms of the following quality factors [5–7]: the current figure of merit,

$$F_i = \frac{p}{c_V}\ [\mathrm{m\ V^{-1}}];$$

the voltage figure of merit,

$$F_V = \frac{p}{c_V\varepsilon\varepsilon_0}\ [\mathrm{V\ m\ J^{-1}}]; \qquad (3)$$

and the noise figure of merit,

$$F_D = \frac{p}{c_V\sqrt{\varepsilon\varepsilon_0\tan\delta}}\ [\mathrm{V\ m^{3/2}\ J^{-1/2}}],\ [\mathrm{Pa^{-1/2}}].$$

Here, c_V is the heat capacity at constant volume, ε is the permittivity, and $\tan\delta$ is the dielectric loss tangent of the material.

Thus, for the comparative characterization of the material with its application in IR detectors in mind, we measured its pyroelectric constant, permittivity, and dielectric loss tangent over a wide range of temperatures and applied dc voltages. Properly adjusting the phase relations between the source signal and pyroelectric current, we were able to detect the latter on the background of parasitics.

The samples of the PMN–PT ferroelectric ceramics were 400 μm thick and had a diameter of 8 mm. The electrodes were made by firing-in of silver. As the source of the modulated radiation, we used an IRS-2-870-6 light-emitting diode (100 mW at 0.87 μm).

The pyroelectric constant was measured in the temperature range from 278 to 350 K, which is of most interest for applications. The temperature was kept to within ±0.1 K. The sample temperature was controlled with the help of a Peltier element coupled with a temperature-stabilization circuit. The pyroelectric constant was calculated from the measured pyroelectric current amplitude by formula (2). In a single measurement, the pyroelectric constant was calculated with an accuracy of $\Delta p \cong 10^{-2}$ [10^{-4} C/(m^2 K)] [15].

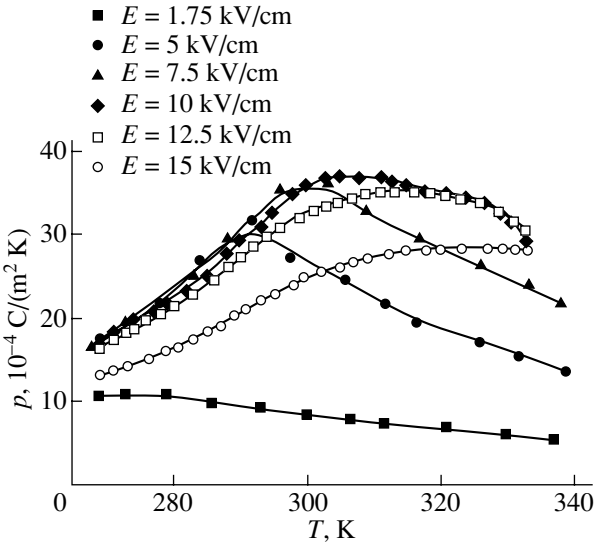


Fig. 1. Induced pyroelectric constant of PMN–PT versus temperature for various biasing fields.

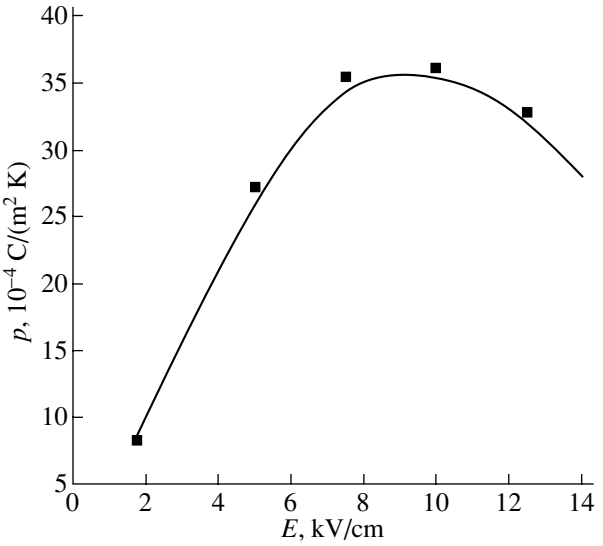


Fig. 2. Induced pyroelectric constant of PMN–PT ceramics versus electric field intensity at 300 K.

Taking the temperature dependences of the induced pyroelectric current (Fig. 1), we calculated the pyroelectric constant in PMN–PT ferroelectric ceramics at various biasing fields. Figure 2 plots the pyroelectric constant as a function of the applied field at 300 K. From these dependences, we found the optimum ranges of temperature (293–325 K) and electric field intensity (8–12 kV/cm) where the pyroelectric constant achieves a maximum (up to $p = 38 \times 10^{-4}$ C/(m^2 K)). In this temperature range, we also measured the relative permittivity and dielectric loss tangent of the samples, using the standard technique [14]. These measurements allowed us to calculate the figures of merit of the material from expressions (3) and plot them versus temperature and external electric field intensity (Figs. 3, 4).

DETECTIVITY OF PYROELECTRIC MATERIALS AND COMPARATIVE CHARACTERIZATION OF UNCOOLED MID-IR PHOTODETECTORS

Table 1 summarizes the highest values of the pyroelectric constant found experimentally and the values of other physical parameters of the material. From these values, we calculated the figures of merit of PMN–PT ferroelectric ceramics at 300 K. This table also lists the same characteristics published earlier for other conventional pyroelectric materials, as well as for the materials used in dielectric bolometers. It should be noted that the data given in Table 1, especially those for dielectric bolometers, may only be used for qualitative comparison, since the published values of the pyroelectric constants of the materials listed are highly spread. For

Table 1. Physical parameters and figures of merit for a number of pyroelectric materials at 300 K

	Conventional pyroelectric materials			Materials for dielectric bolometers		
	TGS	PZT	PVDF	BST, 1 V/μm	PST, 15–35 V/μm	PMN–PT, 1 V/μm
p, 10^{-4} C/(m^2 K)	2.8	3.8–4	0.25	8	1.5–4.5…55	38
ε	38	389–558	9	4000	10000	6000
tan δ	0.01	0.012	0.03	0.005	<0.007	0.002
c_v, 10^6 J/(m^3 K)	2.3	2.5	2.3	2.5	2.7	2.4
F_V, (V m^2)/J	0.36	0.028–0.038	0.14	0.01	0.025	0.06
F_i, 10^{12} m/V	120	–	11	320	2000	1500
F_D, 10^{-5} (V m$^{3/2}$)/J$^{1/2}$ (Pa$^{-1/2}$)	66	18.7–20.3	7	4	1.6–3.6	6
References	[9]	[12]	[9]	[6]	[8, 10]	This work

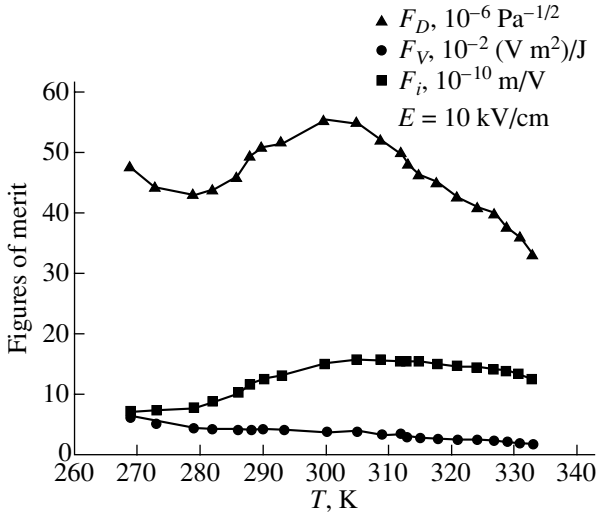


Fig. 3. Figures of merit of PMN–PT relaxer ceramics versus temperature.

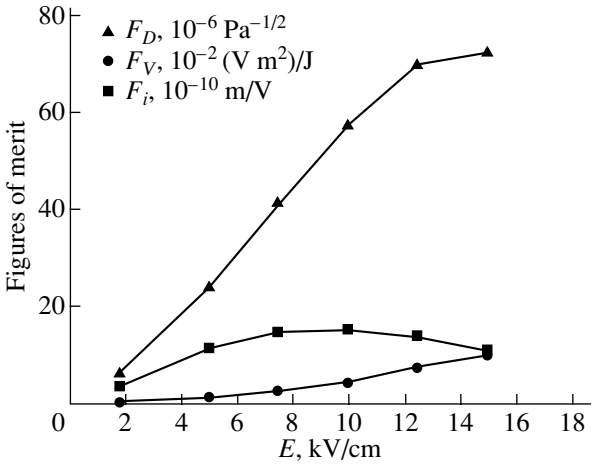


Fig. 4. Figures of merit of PMN–PT relaxer ceramics versus electric field at 300 K.

example, for $Pb(ScTa)_{0.5}O_3$ (PST) ceramics, the spread in the pyroelectric constant exceeds one order of magnitude: from 55×10^{-4} [8] to 1.5×10^{-4} C/(m^2 K) [10]. In our opinion, the basic reason is incorrect measurements in the presence of an external biasing field: the pyroelectric current in the samples and, hence, the pyroelectric constant are usually overestimated. Nevertheless, Table 1 shows that PMN–PT ceramics is comparable to, or even surpasses, the other materials, except for TGS, in performance. PMN–PT has virtually the same characteristics as PST, which also is a relaxer, in external fields that are 15 to 30 times weaker. Note that PMN–PT ceramics maintains high figures of merit in a wide temperature range of about 50°C, while the working range of $(Ba,Sr)TiO_3$ (BST) and PST is as small as several degrees. Also, PMN–PT is more promising for mass production, because it is much cheaper than PST (PMN–PT is free of relatively expensive scandium) and single-crystal TGS. The data obtained experimentally and those given in Table 1 were used to calculate the current and voltage sensitivities, as well as the detectivity, of photodetectors that may be designed around these materials. The calculation results are summarized in Table 2 along with the same characteristics of the other known detectors (photoresistors, photodiodes, conventional pyroelectric detectors, and dielectric bolometers) studied elsewhere.

As follows from Table 2, even in the mid-IR range, the analytical detectivity of dielectric bolometers is close to that of photoresistors in order of magnitude. Furthermore, the voltage sensitivity of the bolometers is comparable to that of photoresistors and the current sensitivity of the bolometers is higher. Certainly, uncooled InAs photodiodes feature a much higher current sensitivity and detectivity, but their low dark resistance (hundreds of ohms at $\lambda = 3.4$ μm or several ohms

Table 2. Characteristics of uncooled (300 K) mid-IR detectors

	Photoresistors PbSe	Photodiodes InAs	Conventional pyroelectric materials			Dielectric bolometers		
			TGS	PZT (thin film)	PVDF	BST (thin film)	PMN–PT 1 V/μm	
Frequency, Hz	500	500	5	60	5	1000	5	1000
Detector dimensions:								
diameter D, mm	2.6×2.6	1×1	8	2.24×0.36	0.4	0.2×0.2	8	0.2×0.2
thickness d, μm			400		25		400	
Spectral range, μm	2…4.5	3.3 4.3 4.7	Entire range	Entire range	Entire range	Entire range	Entire range	
R_V, V/W	10^3…10^2	200 15 4	1440	–	100	1200	151.2	1500
R_i, 10^{-6} A/W	10^{-2}…10^{-3}	10^6	0.3	–	0.44	–	3.75	–
D*, 10^8 cm/(Hz$^{1/2}$ W)	1…0.2	150…5	15	1.8	0.07	2.9	1.36	4.35
References	[2]	[3]	[1, 9]	[12]	[9]	[7]	This work	

at $\lambda = 4.7$ μm) makes efficient current-to-voltage conversion with a high signal-to-noise ratio difficult. As the radiation wavelength increases, the detectivity of the photodiodes drops and the characteristics of pyroelectric detectors, which are wavelength independent, become unattainable for other types of uncooled detectors.

CONCLUSIONS

By the example of $0.9PbMg_{1/3}Nb_{2/3}O_3$–$0.1PbTiO_3$, a typical relaxer ferroelectric, it is shown that relaxer ferroelectrics are promising materials for uncooled mid-IR detectors, since they combine high figures of merit and wide working temperature ranges. Moreover, the material is fairly cheap and opens the possibility of manufacturing large-area detectors. High values of the pyroelectric constant, $p = 30 \times 10^{-4}$ C/(m^2 K) (the highest value is $p = 38 \times 10^{-4}$ C/(m^2 K)), and high figures of merit of PMN–PT are achieved under external fields of 8 to 12 kV/cm in a wide temperature range from 290 to 320 K. As the radiation wavelength increases, the detectivity of semiconductor photodiodes and photoresistors decreases. At the same time, the characteristics of pyroelectric detectors, being independent of the wavelength, surpass IR detectors of other types in a number of parameters. One more advantage of the detectors based on relaxer ferroelectrics, in which the field-induced pyroelectric effect is observed, is that basically they allow for applied bias modulation, rather than incident light modulation (as in conventional pyroelectric detectors).

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 02-02-16389), a grant from the President of the Russian Federation (grant no. NSh-2168.2003.2), the program "Solid-State Nanostructures," and the Department of Physical Sciences at the Russian Academy of Sciences.

REFERENCES

1. A. van der Ziel, *Noise in Measurements* (Wiley, New York, 1976; Mir, Moscow, 1979).
2. A. O. Olesk, *Photoresistors* (Énergiya, Moscow, 1966) [in Russian].
3. B. A. Matveev, M. Aĭdaraliev, G. A. Gavrilov, *et al.*, Sens. Actuators B **51**, 233 (1998).
4. M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Oxford Univ. Press, Oxford, 1977; Mir, Moscow, 1981).
5. M. Daglish, Integr. Ferroelectr. **22**, 473 (1988).
6. R. W. Whatmore, P. C. Osbond, and N. M. Shorrocks, Ferroelectrics **76**, 351 (1987).
7. M. Noda, K. Inoue, M. Ogura, *et al.*, Sens. Actuators A **97–98**, 329 (2002).
8. N. M. Shorrock, R. W. Whatmore, and P. C. Osbond, Ferroelectrics **106**, 387 (1990).
9. S. Bauer and S. B. Lang, IEEE Trans. Dielectr. Electr. Insul. **3**, 647 (1996).
10. M. A. Todd, P. P. Donohue, M. A. Harper, *et al.*, Integr. Ferroelectr. **35**, 115 (2001).
11. M. Noda, H. Zhu, H. Xu, *et al.*, Integr. Ferroelectr. **35**, 31 (2001).
12. W. Liu, J. S. Ko, and W. Zhu, Integr. Ferroelectr. **35**, 127 (2001).
13. A. P. de Kroon, S. C. Dunn, and R. W. Whatmore, Integr. Ferroelectr. **35**, 209 (2001).
14. E. P. Smirnova, S. E. Aleksandrov, K. A. Sotnikov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **45**, 1245 (2003) [Phys. Solid State **45**, 1305 (2003)].
15. S. E. Aleksandrov, G. A. Gavrilov, A. A. Kapralov, *et al.*, Proc. SPIE **5381**, 128 (2004).

Translated by A. Khzmalyan

Technical Physics, Vol. 49, No. 9, 2004, pp. 1181–1190. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 77–85.

SOLID-STATE ELECTRONICS

Theoretical Concepts of Operation of a Semiconductor Detector Based on a *p–n* Junction

L. A. Bakaleĭnikov, E. Yu. Flegontova, K. Yu. Pogrebitskiĭ, and I. V. Eremin

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: fl.xiees@mail.ioffe.ru

Received February 25, 2004

Abstract—Collected charge originating in a semiconductor detector of the p^+–n–n^+ type as a result of interaction with a monoenergetic electron beam with energies in the range from 7 to 25 keV is calculated. Generation of electron–hole pairs (EHPs) is calculated using the Monte Carlo method. In the context of the diffusion–drift model, an analytic expression for the contribution of generated EHPs to the detected signal is derived. It is shown that the losses of charge to recombination in the course of transport significantly affect the shape of detected signal. The comparison of simulated energy spectra with experimentally measured spectra shows good agreement between theory and experiment. Thus, the basics of a theoretical approach that makes it possible to calculate the operational characteristics of semiconductor detectors are developed; as a result, the parameters of these detectors can be optimized in designing the practically important semiconductor proportional detectors to be used in analytical methods. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Current progress in science and technology requires the design and development of systems for control and evaluation of basic phenomena in the field of detection of various types of radiation. Applications of these efforts are numerous; they range from ecological monitoring of the state of the environment via the control of radiation-hazardous works to the problem of ultrafine tool-related and scientific-analytical instrument making and also of diagnostic medical instrumentation.

An important aspect in the aforementioned problem is exactly the development of detectors of ionizing radiation as the specific product that makes it possible to implement various physical methods for monitoring and diagnostics.

Semiconductor proportional detectors (SPDs) occupy a highly important place among the detectors based on different physical phenomena. This type of detector is most promising (compared to gas-filled detectors, channel multipliers, and scintillation detectors). The advantages of SPDs are following: (i) small size of the active region; (ii) the possibility of integration with the detection circuit on the same chip; and (iii) the possibility of attaining the position sensitivity using segmented detecting elements and planar silicon technology.

In particular, special interest is attracted to the use of SPDs in modern methods of diagnostics of solids, in which case information about the objects under study is gained from the analysis of the distribution of electrons emitted from the sample surface. These methods include EXAFS/XANES spectroscopy [1], Auger spectroscopy and X-ray photoelectron spectroscopy [2], the method of standing X-ray waves [3], and other diagnostic methods [4–7]. However, until now, only the reports concerned with success in the use of SPDs in specific studies have been published [8, 9]; the theory of interaction of electrons with actual semiconductor structure of the microchip itself has hardly been considered at all. Thus, it is evident that the above-formulated object of this study is of current interest.

In this paper, we consider and develop the theoretical concepts of operation of a semiconductor detector of electrons. Using a rigid specification of the SPD structure, we find the methods for optimization of operation of the detectors under consideration. It is shown that practical optimization of SPD parameters depends almost to the same extent on the internal structure of the active component as on the external factors (in particular, on the applied reverse-bias voltage).

FORMULATION OF THE PROBLEM

Operation of semiconductor detectors is based on generation of electron–hole pairs (EHPs) within the p^+–n junction and collection of these pairs under the applied reverse-bias voltage. The energy of detected electrons is found to be related to the amplitude of the electric signal that appears at the detector electrodes; in turn, this circumstance can be used to determine the initial electron spectrum. In this context, the problem of clarifying the dependence of the signal on the detector parameters and the energy of electrons is of current interest.

1063-7842/04/4909-1181$26.00

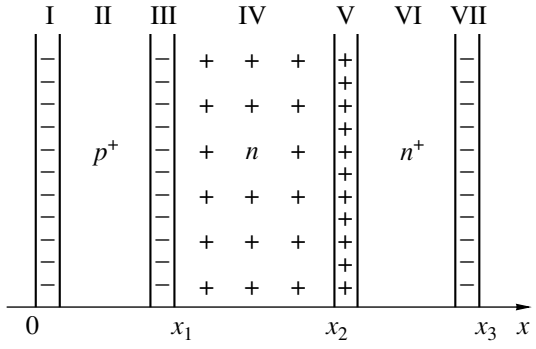


Fig. 1. The structure of a semiconductor electron detector: $x_1 = 0.1$ μm, $x_2 - x_1 = 300$ μm, and $x_3 - x_2 = 0.1$ μm.

In this paper, we pose the problem of numerical simulation of the signal generation by a monoenergetic electron beam in a semiconductor detector. This problem is solved in two stages. In the first stage, the distribution of EHPs generated by an incident electron in the detector is calculated; in the second stage, the collected charge arising in the semiconductor structure under the effect of applied reverse bias is calculated.

The simulated semiconductor detector [10] is a planar p^+–n–n^+ semiconductor structure fabricated on the basis of high-resistivity n-Si (Fig. 1). The thickness of the p^+- and n^+-type contact regions $\Delta = 1000$ Å; the thickness of the n-type region $L = 300$ μm. The n- and n^+-type regions were doped with phosphorus, and the p^+-type region was doped with boron. The doping-impurity concentration was $N_D \approx N_A \sim 10^{19}$ cm^{-3} in the n^+- and p^+-type regions and was $N_D \sim 10^{12}$ cm^{-3} in the n-type region. Under operating conditions, the n-type region is completely depleted of electrons, which corresponds to the reverse voltage $U = 80$ V applied to the p^+–n junction at the aforementioned parameters of the semiconductor structure. In this case, a dark current with a density on the order of 1 nA/cm^2 flows through the structure. The n^+-type contact serves as the entrance window for electrons to be detected in the case chosen for simulation; the electron beam is incident on the sample surface along the normal.

SIMULATION OF GENERATION OF ELECTRON–HOLE PAIRS

In order to calculate the distribution of EHPs in the detector, we need to know the distribution of electrons in the incident beam. This distribution can be found from the kinetic equation

$$\begin{aligned}\Omega \mathrm{grad}_{\mathbf{r}}\Phi(r, \Omega, E) &= \int_{4\pi}\frac{dw_{\mathrm{el}}(E; \Omega', \Omega)}{d\Omega}\Phi(r, \Omega', E)d\Omega' \\ &- \int_{4\pi}\frac{dw_{\mathrm{el}}(E; \Omega, \Omega')}{d\Omega'}d\Omega'\Phi(r, \Omega, E) \\ &+ \int_{4\pi}\int_{0}^{E_0 - E}\frac{d^2w_{\mathrm{in}}(E + Q, \Omega', E, \Omega)}{dQd\Omega}\Phi(r, \Omega', E + Q)d\Omega' dQ \\ &- \int_{4\pi}\int_{0}^{E}\frac{d^2w_{\mathrm{in}}(E, \Omega, E - Q, \Omega')}{dQd\Omega'}d\Omega' dQ\Phi(r, \Omega, E)\,.\end{aligned} \tag{1}$$

Here, $\Phi(r, \Omega, E)$ is the differential flux density of electrons with the energy E and the direction of motion Ω at the point $\mathbf{r}$; and

$$\frac{dw_{\mathrm{el}}(E; \Omega, \Omega')}{d\Omega'}, \quad \frac{d^2w_{\mathrm{in}}(E, \Omega, E', \Omega')}{dQd\Omega'}$$

are the differential reciprocal free paths of electrons with respect to the elastic and inelastic collisions. Equation (1) was solved using the Monte Carlo method. The single-scattering model described previously [11] was used; according to this model, the result of each event of interaction between an electron and the material is controlled by differential cross sections of the elastic and inelastic scattering. In order to determine the results of the elastic scattering, we used the differential Mott cross section. Inelastic interaction of electrons with the material is described by the doubly differentiated reciprocal free path $d^2w_{\mathrm{in}}(E, \Omega, E', \Omega')/dQd\Omega'$; i.e., this interaction is controlled by the probability that an electron with an energy E loses the energy $Q = E - E'$ and is scattered by the angle $\theta = \arccos(\Omega\Omega')$ per unit path length. The doubly differentiated reciprocal free path can be determined from the data on the permittivity of the material [12–14]. Flegontova *et al*. [15] described the procedure for calculation of the above quantity and the algorithm for drawing the results of inelastic scattering; this algorithm makes it possible to reduce appreciably the computation time. The data on differential cross sections of elastic and inelastic scattering calculated by us can be found in the electronic archive [16].

The depth distribution of EHPs generated by the electron beam was derived in several models. In the first (most crude) model, it was assumed that all energy lost in an inelastic collision was spent on the formation of EHPs. In this case, the EHP distribution was calculated using the formula

$$\rho(\mathbf{r}) = \int_{4\pi}\int_{0}^{E_0}\beta(E)\Phi(r, \Omega, E)dEd\Omega/\Delta E_{\mathrm{eh}}, \tag{2}$$

where

$$\beta(E) = \int_{4\pi}\int_{0}^{E} \frac{d^2 w_{in}(E, \Omega, E-Q, \Omega')}{dQd\Omega'} QdQd\Omega'$$

and ΔE_{eh} is the average EHP-formation energy.

In the second model, the transport of secondary electrons generated in the course of inelastic scattering was taken into account. The trajectory of the entire cascade of secondary electrons was traced for as long as their energy exceeded the given cutoff energy E_c. It was also assumed that, at the point of the trajectory cutoff, the remaining electron energy is completely spent on the formation of EHPs with the average generation energy equal to ΔE_{eh}.

In the third model (as in the second), the trajectories of all electrons were traced; however, it was assumed that each inelastic-collision event led to the formation of a single EHP at the interaction point.

The depth distribution of the generated EHP density calculated according to different models is shown in Fig. 2 for incident-electron energies E_0 = 7 and 20 keV and the cutoff energy E_c = 10 eV. It can be clearly seen that all aforementioned approaches yield the identical shape of the depth distribution of generated EHPs.

This circumstance means that the detailed allowance for the secondary-electron cascade and the transport of the energy transferred to these electrons is not necessary in calculation of the distribution of generated EHPs. In turn, this consideration justifies the use of the first model in calculations, which makes it possible to reduce appreciably the computation time.

In addition to the calculation of the EHP depth distribution, it is also of interest to simulate the distribution of electrons in the lost energy, since the corresponding dependence is related to the experimentally studied characteristic of the detector, i.e., the amplitude distribution of the voltage pulses equivalent to the collected charge. It is well known that this distribution is equivalent to the energy spectrum of detected electrons in the situation where the collected charge is proportional to the energy of detected particle.

Assuming that all energy lost by an electron is spent on the formation of EHPs and all EHPs generated in the space-charge region (SCR) contribute to the signal, we can simulate the observed distribution. In order to compare the results of simulation with experimental data, we convolved the calculated dependences with the measured instrument-related spread function for the measurement circuit. The theoretical and experimental curves are shown in Fig. 3 for several electron-beam energies. It can be seen that the calculated and experimental dependences practically coincide at high electron energies; however, a decrease in the energy leads to a discrepancy in the positions of the peaks in the distributions. This behavior is related to the inadequacy of taking into account the contribution of EHPs generated by the electron beam to the signal. In order to take into account this contribution in more detail, we must consider the transport of charge carriers in the detector.

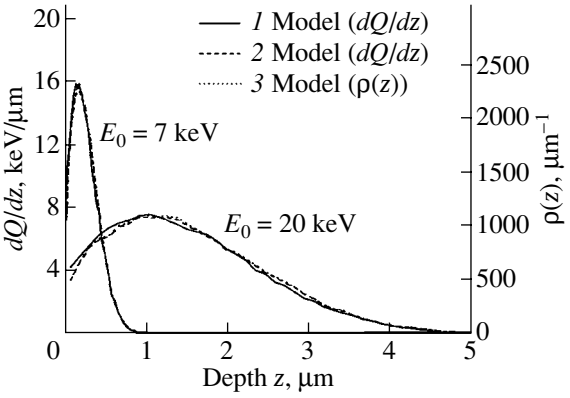


Fig. 2. The depth distribution of the energy dissipated by electrons dQ/dz and the density of generation of electron–hole pairs $\rho(z)$ for electron-beam energies of 7 and 20 keV.

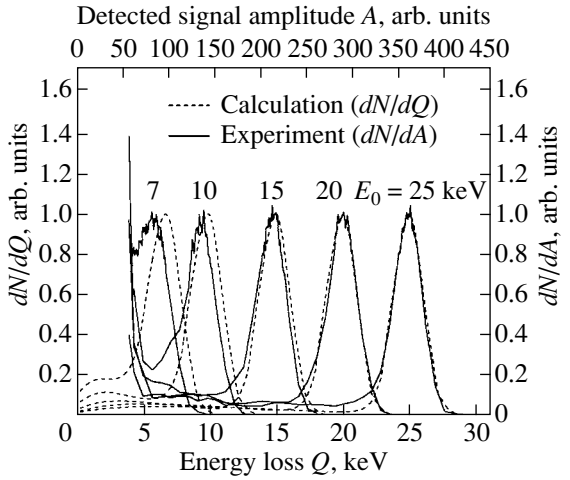


Fig. 3. Comparison of experimental signal spectrum for a detector irradiated with monoenergetic electrons with energies in the range from 7 to 25 keV with the results of simulation taking into account the contribution of electron–hole pairs generated within the space-charge region to the signal.

SIMULATION OF THE CHARGE-CARRIER TRANSPORT IN THE DETECTOR

The contribution of an EHP generated at a certain depth to the detected signal is obtained by solving a system of equations that describe the transport of charge carriers in the system under consideration in the diffusion–drift approximation. This system includes the Poisson equation for an electric field, continuity equations for concentrations of electrons and holes, and expressions for the currents and recombination of charge carriers.

We assume that the coordinate axis is normal to the detector surface and the coordinate origin $x = 0$ is located at the interface between the p^+-type region and the metallic electrode. We designate the coordinate of the interface between the p^+- and n-type regions by x_1, the coordinate of the interface between the n- and n^+-type regions by x_2, and the coordinate of the interface between the n^+-type region and the metallic electrode by x_3 (see Fig. 1). The distribution of concentrations for electrons $n(x)$ and holes $p(x)$ is determined from the continuity equations

$$\frac{1}{e}\frac{\partial j_p}{\partial x} + R = g, \tag{3}$$

$$\frac{1}{e}\frac{\partial j_n}{\partial x} + R = -g, \tag{4}$$

where the electron and hole fluxes have the conventional form in the diffusion–drift approximation, i.e.,

$$\frac{1}{e}j_p = -\mu_p\frac{d\varphi}{dx}p - D_p\frac{\partial p}{\partial x}, \tag{5}$$

$$\frac{1}{e}j_n = -\mu_n\frac{d\varphi}{dx}n + D_n\frac{\partial n}{\partial x}. \tag{6}$$

Here, μ_n, μ_p, D_n, and D_p are the mobilities and diffusion coefficients of electrons and holes. The recombination rate of nonequilibrium charge carriers via a single recombination level is given by the following expression in the Shockley–Read–Hall approximation [17]:

$$R = \frac{(pn - p_0n_0)}{\tau_{0p}(n + n_1) + \tau_{0n}(p + p_1)}. \tag{7}$$

Here, p_0 and n_0 are the thermodynamically equilibrium concentrations of electrons and holes; $\tau_{0n} = 1/\alpha_nN_t$; $\tau_{0p} = 1/\alpha_pN_t$; L_n, L_p are the coefficients of electron and hole capture at the impurity level; N_t is the impurity concentration; $n_1 = n_0(1 - f_0)/f_0$; $p_1 = p_0f_0(1 - f_0)/f_0$; and f_0 is the equilibrium probability of filling the impurity level with electrons.

We assume that the rate of generation g of electrons and holes by external radiation is equal to

$$g = G\delta(x - x_0), \tag{8}$$

where G is the density of generation of charge carriers in the plane $x = x_0$ (G has the dimension of flux); i.e., we estimate the contribution of a stationary point source located at a depth x_0 to the signal.

The potential distribution $\varphi(x)$ is described by the Poisson equation

$$-\frac{d^2\varphi}{dx^2} = \frac{4\pi e}{\varepsilon}(p + p_t - n - n_t), \tag{9}$$

where p_t is the concentration of positively charged donors (bound holes) and n_t is the concentration of negatively charged acceptors.

In the system under consideration, $n_t \approx N_A \sim 10^{19}$ cm^{-3} and $p_t \approx 0$ in the p^+-type region, $p_t \approx N_D \sim 10^{12}$ cm^{-3} and $n_t \approx 0$ in the n-type region, and $p_t \approx N_D \sim 10^{19}$ cm^{-3} and $n_t \approx 0$ in the n^+-type region.

The system of Eqs. (3)–(9) under consideration is a sixth order system and requires six boundary conditions for its solution. We set the fluxes of nonequilibrium charge carriers equal to zero at the metal–semiconductor interfaces $x = 0$ and x_3 (the interfaces between the detector and metal electrodes); i.e.,

$$j_p(0) = 0, \quad j_n(x_3) = 0. \tag{10}$$

We assume that the fluxes of majority charge carriers are small compared to those of thermal emission; as a result, we have

$$p(0) = p_s, \quad n(x_3) = n_s, \tag{11}$$

where p_s and n_s are the equilibrium concentrations of charge carriers at the interfaces with metal electrodes; these concentrations are related to the contact potential differences U_1 and U_4 at the boundaries $x = 0$ and x_3 by the expressions [18]

$$p_s = p_0\exp(U_1/kT), \quad n_s = n_0\exp(-U_4/kT). \tag{12}$$

The following boundary conditions should be satisfied for the potential:

$$\varphi(0) = 0, \quad \varphi(x_3) = U + \sum_{k=1}^{4} U_k. \tag{13}$$

Here, U is the applied external voltage and U_k are the contact potential differences at the regions' boundaries.

We require the continuity of potential, field strength, concentrations, and currents at the internal regions' boundaries $x = x_1$ and x_2.

In order to solve the system of Eqs. (3)–(13), we employ the conventional approach [19]. We divide the semiconductor into the SCRs (I, III, IV, V, and VII) and the quasi-neutral regions (II and VI) (see Fig. 1). It is taken into account in this separation that the n-type region is completely depleted of electrons if a reverse external bias voltage $U = 80$ V is applied.

Assuming that the external generation of charge carriers is insignificant and the effect of current on the potential distribution can be disregarded, we determine the potential in the approximation of the depleted (enriched) layer; i.e.,

in region I ($0 < x \le d_0^+$), we have

$$\Psi(x) = \frac{1}{2L_p^2}(x - d_0^+) + \frac{eU_1}{kT}, \quad d_0^+ = L_p\sqrt{-2U_1e/kT}; \tag{14}$$

in region II ($d_0^+ < x \le x_1 - d_0^-$),

$$\Psi(x) = \Psi(d_0^+) = eU_1/kT; \tag{15}$$

in region III ($x_1 - d_0^- < x \le x_1$),

$$\Psi(x) = \frac{1}{2L_p^2}(x - (x_1 - d_1^-))^2 + eU_1/kT,$$

$$d_1^- = \frac{(x_2 - x_1)}{1 + p_0/n_0^+}\left\{\frac{n_0}{n_0^+} - 1 + \sqrt{\left(\frac{n_0}{n_0^+} - 1\right)^2 - \left(\frac{1}{p_0} + \frac{1}{n_0^+}\right)\left(\frac{n_0^2}{n_0^+} - n_0 - \frac{\varepsilon}{2\pi e}\frac{(U + U_2 + U_3)}{(x_2 - x_1)^2}\right)}\right\}; \quad (16)$$

in region IV ($x_1 < x < x_2$),

$$\Psi(x) = -\frac{1}{2L_n^2}(x - x_2)^2 + \frac{1}{L_{n^+}^2}d_2^+(x - x_2) - \frac{1}{2L_{n^+}^2}(d_2^+)^2 + e\left(U + \sum_{k=1}^{3} U_k\right)/kT; \quad (17)$$

in region V ($x_2 \le x < x_2 + d_2^+$),

$$\Psi(x) = -\frac{1}{2L_{n^+}^2}(x - (x_2 + d_2^+))^2 + e\left(U + \sum_{k=1}^{3} U_k\right)/kT, \quad (18)$$

$$d_2^+ = \frac{p_0}{n_0^+}d_1^- - \frac{n_0}{n_0^+}(x_2 - x_1);$$

in region VI ($x_2 + d_2^+ \le x < x_3 - d_3^-$),

$$\Psi(x) = e\left(U + \sum_{k=1}^{3} U_k\right)/kT; \quad (19)$$

and in region VII ($x_3 - d_3^- \le x < x_3$),

$$\Psi(x) = -2\ln\left(\cos\left(\frac{1}{\sqrt{2}L_{n^+}}(x - (x_3 - d_3^-))\right)\right) + e\left(U + \sum_{k=1}^{3} U_k\right)/kT, \quad (20)$$

$$d_3^- = \sqrt{2}L_{n^+}\arccos(\exp(-eU_4/2kT)).$$

Here, $\Psi(x)$ is the dimensionless potential related to $\varphi(x)$ by the expression

$$\Psi(x) = e\varphi(x)/kT, \quad L_p = \left(\frac{\varepsilon kT}{4\pi e^2 p_0}\right)^{1/2}, \quad L_n = \left(\frac{\varepsilon kT}{4\pi e^2 n_0}\right)^{1/2}, \quad L_{n^+} = \left(\frac{\varepsilon kT}{4\pi e^2 n_0^+}\right)^{1/2}.$$

It was assumed in deriving these expressions that the contact potential difference at the interfaces with metal electrodes are $U_1 < 0$ and $U_4 < 0$. Estimations show that these assumptions do not introduce any significant errors into the final result. It is noteworthy that, in the remaining system for the concentrations and currents of charge carriers in the space-charge layers, we can disregard the volume recombination owing to a short transit time of charge carriers compared to the corresponding lifetime.

We can use the following approximation for the recombination term in quasi-neutrality region II:

$$R = \frac{n - n_0}{\tau_n}. \quad (21)$$

Here, τ_n is the lifetime of nonequilibrium electrons and n_0 is the equilibrium electron concentration.

A similar situation also takes place in the quasi-neutrality region VI. In this case, the recombination term is approximated by the expression

$$R = \frac{p - p_0}{\tau_p}. \quad (22)$$

Thus, the system of equations for determination of currents and concentrations is found to be linear for all regions.

In what follows, we are most interested in the signal that appears in the detector as a result of generation of EHPs; i.e., we are interested in the difference between the currents through the detector in the presence and in the absence of generation. Since the system of equations for currents and concentrations is linear, the difference between solutions for $G \neq 0$ and $G = 0$ satisfies the system of Eqs. (3)–(6) with the recombination term

$$R = \begin{cases} 0 & x \in [d_0^+, x_1 - d_1^-] \cup [x_2 + d_2^+, x_3 - d_3^-] \\ n/\tau_n & x \in [d_0^+, x_1 - d_1^-] \\ p/\tau_p & x \in [x_2 + d_2^+, x_3 - d_3^-] \end{cases} \quad (23)$$

and homogeneous boundary conditions

$$j_p(0) = 0, \quad j_n(x_3) = 0, \quad p(0) = 0, \quad n(x_3) = 0. \quad (24)$$

In order to solve this system, we note that, in the SCRs, the currents $j_n(x)$ and $j_p(x)$ remain constant. In the plane $x = x_0$, the generation conditions (8) give rise to a jump in the currents

$$j_p|_{x_0+0} - j_p|_{x_0-0} = eG, \quad j_n|_{x_0+0} - j_n|_{x_0-0} = -eG. \quad (25)$$

In this case, the distribution of currents can be easily found by integration of Eqs. (5) and (6); as a result, we

obtain

$$n(x) = \left(n(x_b^i)\exp(-\Psi(x_b^i)) + \frac{j_n^i}{eD_n}\int_{x_b^i}^{x}\exp(-\Psi(\xi))d\xi\right)\exp(\Psi(x)), \quad (26)$$

$$p(x) = \left(p(x_b^i)\exp(\Psi(x_b^i)) - \frac{j_p^i}{eD_p}\int_{x_b^i}^{x}\exp(\Psi(\xi))d\xi\right)\exp(-\Psi(x)). \quad (27)$$

Here, x varies within each of separate SCRs; x_b^i is the boundary point of a specific SCR; and j_n^i and j_p^i are the currents in this region. For the quasi-neutrality region II, the system of equations for the concentration n and currents j_n of minority charge carriers can be written as

$$\frac{dn}{dx} = \frac{1}{eD_n}j_n, \quad \frac{dj_n}{dx} = e\frac{n}{\tau_n}. \quad (28)$$

By integrating this system, we obtain the expressions for the concentration and currents of minority charge carriers in region II; i.e.,

$$n(x) = C_1\exp\left(\frac{x}{\sqrt{\tau_n D_n}}\right) + C_2\exp\left(-\frac{x}{\sqrt{\tau_n D_n}}\right), \quad (29)$$

$$j_n(x) = e\sqrt{\frac{D_n}{\tau_n}}\left(C_1\exp\left(\frac{x}{\sqrt{\tau_n D_n}}\right) - C_2\exp\left(-\frac{x}{\sqrt{\tau_n D_n}}\right)\right). \quad (30)$$

The current of majority charge carriers in this region can be found by integration of Eq. (3); we obtain

$$j_p(x) = -e\sqrt{\frac{D_n}{\tau_n}}\left(C_1\exp\left(\frac{x}{\sqrt{\tau_n D_n}}\right) - C_2\exp\left(-\frac{x}{\sqrt{\tau_n D_n}}\right)\right) + C_3. \quad (31)$$

Taking into account that $j_p(x)$ and $p(x)$ are related by Eq. (5), we obtain

$$p(x) = \frac{D_n}{D_p}\left(C_1\exp\left(\frac{x}{\sqrt{\tau_n D_n}}\right) + C_2\exp\left(-\frac{x}{\sqrt{\tau_n D_n}}\right)\right) + C_3 x + C_4. \quad (32)$$

Similar results are also valid for the quasi-neutrality region VI, which makes it possible to derive the expressions for the concentration and current of holes in forms similar to those in formulas (29) and (30). In this case, the concentration and current of electrons are determined from Eqs. (4) and (6); the resulting expressions are similar to those in (31) and (32).

The final form of the dependence of charge-carrier concentrations on x in various SCRs can be obtained by substituting the potentials (14)–(20) into expressions (26) and (27). Thus, for example, we have the following expressions for region I:

$$n(x) = \left(n(d_0^+) + \frac{\sqrt{2}L_p j_n(d_0^+)}{eD_n}F\left(\frac{(x-d_0^+)}{\sqrt{2}L_p}\right)\right) \times \exp\left(\frac{(x-d_0^+)^2}{2L_p^2}\right), \quad (33)$$

$$p(x) = \left(p(d_0^+) - \frac{\sqrt{2}L_p j_p(d_0^+)}{eD_p}G\left(\frac{(x-d_0^+)}{\sqrt{2}L_p}\right)\right) \times \exp\left(-\frac{(x-d_0^+)^2}{2L_p^2}\right). \quad (34)$$

Here,

$$G(x) = \int_0^x \exp(\tau^2)d\tau, \quad F(x) = \int_0^x \exp(-\tau^2)d\tau. \quad (35)$$

In regions III, IV, and V, the solution has a structure similar to that in (33) and (34). Calculations yield the following expressions for region VII:

$$n(x) = \left(n(x_3-d_3^-) + \frac{j_n(x_3-d_3^-)}{eD_n}\left(\frac{x-(x_3-d_3^-)}{2} + \frac{L_{n^+}}{2\sqrt{2}}\sin\left(\sqrt{2}\frac{x-(x_3-d_3^-)}{L_{n^+}}\right)\right)\right)\cos^{-2}\left(\frac{x-(x_3-d_3^-)}{\sqrt{2}L_{n^+}}\right), \quad (36)$$

$$p(x) = \left(p(x_3-d_3^-) - \frac{\sqrt{2}L_{n^+}j_p(x_3-d_3^-)}{eD_p} \times \tan\left(\frac{x-(x_3-d_3^-)}{\sqrt{2}L_{n^+}}\right)\right)\cos^2\left(\frac{x-(x_3-d_3^-)}{\sqrt{2}L_{n^+}}\right). \quad (37)$$

In order to determine the constants that appear in the solution, we have to use the additional conditions (24) and (25) and the conditions for continuity of currents and concentrations at the interfaces between the regions. In turn, this procedure makes it possible to determine the total current through the detector j =

$j_p(0) = j_n(x_3)$; this current is caused by external generation and represents the sought contribution of charge carriers generated at $x = x_0$ to the measured signal.

Calculations show that, if the generation point is located in the n-type region ($x_1 < x_0 < x_2$), the total current of EHPs generated by external radiation is given by

$$j(x_0) = -eG\frac{1 - S_1(x_0)/S_2 - S_3(x_0)/S_4}{1 - S_5/S_2 - S_6/S_4}, \qquad (38)$$

where

$$S_1(x_0) = \left(\sqrt{\tau_p D_p}\frac{\cosh(\Delta_2) - 1}{\sinh(\Delta_2)} + \sqrt{2}L_{n^+}G(\delta_3)\right) \times \exp(\delta_4^2 - \delta_1^2 - \delta_2^2 - \delta_3^2)$$

$$- \sqrt{2}L_n\left(G(\delta_4) - G\left(\frac{1}{\sqrt{2}L_n}\left(x_2 - x_0 + \frac{n_0^+}{n_0}d_2^+\right)\right)\right) \times \exp(-\delta_1^2 - \delta_2^2);$$

$$S_2(x_0) = \sqrt{\tau_n D_n}\coth(\Delta_1) + \left(\sqrt{\tau_p D_p}\frac{\cosh(\Delta_2) - 1}{\sinh(\Delta_2)} + \sqrt{2}L_{n^+}G(\delta_3)\right)$$

$$\times \exp(\delta_4^2 - \delta_1^2 - \delta_2^2 - \delta_3^2) - \sqrt{2}L_n(G(\delta_4) - G(\delta_2)) \times \exp(-\delta_1^2 - \delta_2^2) + \sqrt{2}L_p F(\delta_1);$$

$$S_3(x_0) = \left(\sqrt{\tau_n D_n}\frac{\cosh(\Delta_1) - 1}{\sinh(\Delta_1)} + \sqrt{2}L_p G(\delta_1)\right) \times \exp(\delta_4^2 - \delta_1^2 - \delta_2^2 - \delta_3^2) + \sqrt{2}L_n\left(F(\delta_2) - F\left(\sqrt{2}L_n\left(x_2 - x_0 + \frac{n_0^+}{n_0}d_2^+\right)\right)\right)\exp(\delta_4^2 - \delta_3^2);$$

$$S_4 = \sqrt{\tau_p D_p}\coth((\Delta_2)) + \left(\sqrt{\tau_n D_n}\frac{\cosh(\Delta_1) - 1}{\sinh(\Delta_1)} - \sqrt{2}L_p G(\delta_1)\right) \times \exp(\delta_4^2 - \delta_1^2 - \delta_2^2 - \delta_3^2) + \sqrt{2}L_n(F(\delta_2) - F(\delta_4)) \times \exp(\delta_4^2 - \delta_3^2) + \sqrt{2}L_{n^+}F(\delta_3);$$

$$S_5 = \left(\sqrt{\tau_p D_p}\frac{\cosh(\Delta_2) - 1}{\sinh(\Delta_2)} - \frac{d_3^-}{2} - \frac{L_{n^+}}{2\sqrt{2}}\sin\frac{\sqrt{2}d_3^-}{L_{n^+}} - (x_3 - d_3^- - x_2 - d_2^+)\right)\exp(\delta_4^2 - \delta_1^2 - \delta_2^2 - \delta_3^2);$$

$$S_6 = \left(\sqrt{\tau_n D_n}\frac{\cosh(\Delta_1) - 1}{\sinh(\Delta_1)} - \sqrt{2}L_p G\left(\frac{d_0^+}{\sqrt{2}L_p}\right)\exp\left(-\frac{d_0^{+2}}{2L_p^2}\right) - (x_1 - d_1^- - d_0^+)\right)\exp(\delta_4^2 - \delta_1^2 - \delta_2^2 - \delta_3^2);$$

$$\Delta_1 = \frac{(x_1 - d_1^- - d_0^+)}{\sqrt{\tau_n D_n}}; \quad \Delta_2 = \frac{(x_3 - d_3^- - x_2 - d_2^+)}{\sqrt{\tau_p D_p}};$$

$$\delta_1 = \frac{d_1^-}{\sqrt{2}L_p}; \quad \delta_2 = \frac{1}{\sqrt{2}L_n}\left(x_2 - x_1 + \frac{n_0^+}{n_0}d_2^+\right);$$

$$\delta_3 = \frac{d_2^+}{\sqrt{2}L_{n^+}}; \quad \delta_4 = \frac{n_0^+}{n_0}\frac{d_2^+}{\sqrt{2}L_n}.$$

If the generation point is located within the n^+-type region ($x_2 + d_2^+ < x_0 < x_3 - d_3^-$), the total current through the detector is given by

$$j(x_0) = eG\frac{S_7(x_0)/S_2 - S_8(x_0)/S_4}{1 - S_5/S_2 - S_6/S_4}, \qquad (39)$$

where

$$S_7(x_0) = \sqrt{\tau_p D_p}\left(\frac{\cosh(\Delta_2) - 1}{\sinh(\Delta_2)}\coth\left(\frac{x_0 - x_2 - d_2^+}{\sqrt{\tau_p D_p}}\right) - \sinh\left(\frac{x_0 - x_2 - d_2^+}{\sqrt{\tau_p D_p}}\right)\right)\exp(\delta_4^2 - \delta_1^2 - \delta_2^2 - \delta_3^2);$$

$$S_8(x_0) = \sqrt{\tau_p D_p}\left[\coth(\Delta_2)\cosh\left(\frac{x_0 - x_2 - d_2^+}{\sqrt{\tau_p D_p}}\right) - \sinh\left(\frac{x_0 - x_2 - d_2^+}{\sqrt{\tau_p D_p}}\right)\right].$$

Due to a small thickness of SCRs V and VII, the contribution of the sources at $x_2 \le x_0 < x_2 + d_2^+$ and $x_3 - d_3^- \le x < x_3$ to the total signal is found to be negligible, so that $j(x_0)$ was not calculated in these cases.

We can appreciably simplify expressions (38) and (39) by using the relations between numerical values of parameters of the problem. For concentrations $n_0 \sim 10^{11}$–10^{12} cm^{-3} and a thickness of the n-type region on the order of 100 μm, we have $\delta_2^2 - \delta_1^2 = \exp(-(x_2 - x_1)^2/\sqrt{2}L_n^2) \ll 1$, $\sqrt{2}L_{n^+} \gg 1$, and $\sqrt{2}L_p \gg 1$; therefore, in the case of the source located in the depletion layer,

we obtain to a high accuracy

$$j(x_0) = eG(1 - \tilde{S}_1(x_0)/\tilde{S}_2 - \tilde{S}_3(x_0)/\tilde{S}_4), \qquad (40)$$

where

$$\tilde{S}_1(x_0) = \sqrt{2}L_n\left(G\left(\frac{1}{\sqrt{2}L_n}\left(x_2 - x_0 + \frac{n_0^+}{n_0}d_2^+\right)\right) - G(\delta_4)\right)$$
$$\times \exp(-\delta_1^2 - \delta_2^2);$$

$$\tilde{S}_2 = \sqrt{\tau_n D_n}\coth(\Delta_1)$$
$$+ \sqrt{2}L_n(G(\delta_2) - G(\delta_4))\exp(-\delta_1^2 - \delta_2^2);$$

$$\tilde{S}_3(x_0) = \sqrt{2}L_n\left(F(\delta_2) - F\left(\sqrt{2}L_n\left(x_2 - x_0 + \frac{n_0^+}{n_0}d_2^+\right)\right)\right)$$
$$\times \exp(\delta_4^2 - \delta_3^2);$$

$$\tilde{S}_4 = \sqrt{\tau_p D_p}\coth(\Delta_2)$$
$$+ \sqrt{2}L_n(F(\delta_2) - F(\delta_4))\exp(\delta_4^2 - \delta_3^2).$$

If the detector is irradiated with electrons through the n^+-type contact, i.e., if the source is located in the vicinity of x_2, we have $\tilde{S}_1(x_0) = 0$ and

$$j(x_0) = -eG(1 - \tilde{S}_3(x_0)/\tilde{S}_4).$$

Similarly, if the source is located within the quasi-neutral n^+-type region, we obtain

$$j(x_0) = -eGS_8(x_0)/\tilde{S}_2. \qquad (41)$$

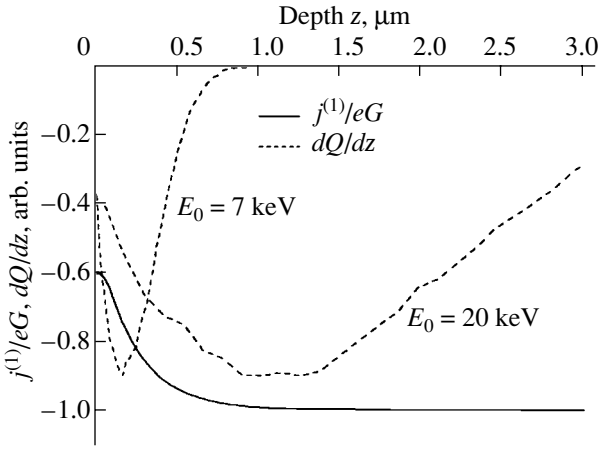


Fig. 4. The dependence of the total current $j(z)$ on the depth of generation of electron–hole pairs z and the distribution of the density of generation of electron–hole pairs over the depth. $p_0 = n_0^+ = 10^{18}$ cm^{-3}, $n_0 = 7 \times 10^{11}$ cm^{-3}, and $\sqrt{D_p\tau_p} = 2 \times 10^{-5}$ cm.

Since $n_0/n_0^+ \ll 1$ and $n_0/p_0 \ll 1$, we can derive the following expressions with an accuracy to the second-order terms:

$$d_1^- = \frac{\varepsilon}{4\pi e}\frac{(U + U_2 + U_3)}{(x_2 - x_1)p_0} + \frac{n_0}{2p_0}(x_2 - x_1),$$

$$d_2^+ = \frac{\varepsilon}{4\pi e}\frac{(U + U_2 + U_3)}{(x_2 - x_1)p_0} - \frac{n_0}{2p_0}(x_2 - x_1).$$

Taking into account that $L_n \sim 1/\sqrt{n_0}$, $L_p \sim 1/\sqrt{p_0}$, and $L_{n^+} \sim 1/\sqrt{n_0^+}$, we can easily analyze the dependence of current on equilibrium concentrations of charge carriers in various regions. The ratio $\tilde{S}_3(x_0)/\tilde{S}_4$ increases with increasing n_0 and decreases with increasing n_0^+/p_0. Naturally, the value of $\tilde{S}_3(x_0)/\tilde{S}_4$ is affected profoundly by the parameter $\sqrt{D_p\tau_p}$ that characterizes the losses to recombination in the quasi-neutral n^+-type region. An increase in $\sqrt{D_p\tau_p}$ leads to a decrease in $\tilde{S}_3(x_0)/\tilde{S}_4$.

Qualitatively, the dependence of current on the parameters of the problem can be interpreted in the following way. Thin regions of a strong screening field in the p^+- and n^+-type regions in the vicinity of the boundaries x_1 and x_2 impede the diffusion of nonequilibrium charge carriers into the quasi-neutral regions where recombination occurs. If the source is located in the vicinity of the boundary x_2, the current is affected only by the thickness d_2^+ of the screening region, the potential-barrier height $U(x_2 + d_2^+) - U(x_2)$, and the diffusion length of nonequilibrium holes in the n^+-type region $\sqrt{D_p\tau_p}$. As n_0 and p_0 increase, the thickness of the screening region decreases, whereas this thickness is nearly independent of n_0^+ (as long as the condition $n_0/n_0^+ \ll 1$ is satisfied). However, as n_0^+/p_0 increases, the height of the potential barrier that inhibits the diffusion of holes into the n^+-type region increases. Therefore, an increase in n_0 and p_0 and a decrease in n_0^+ lead to an increase in the characteristic generation depth within which the effect of losses to recombination is appreciable. It is worth noting that a variation in p_0 and n_0^+ affects only slightly the current if the ratio n_0^+/p_0 is retained constant.

In our calculations, the parameters of the problem were varied within approximately an order of magnitude around the values of $n_0^+ \sim 10^{19}$ cm^{-3}, $p_0 \sim 10^{19}$ cm^{-3}, $n_0 \sim 10^{12}$ cm^{-3}, and $\sqrt{D_p\tau_p} \sim 10^{-5}$ cm.

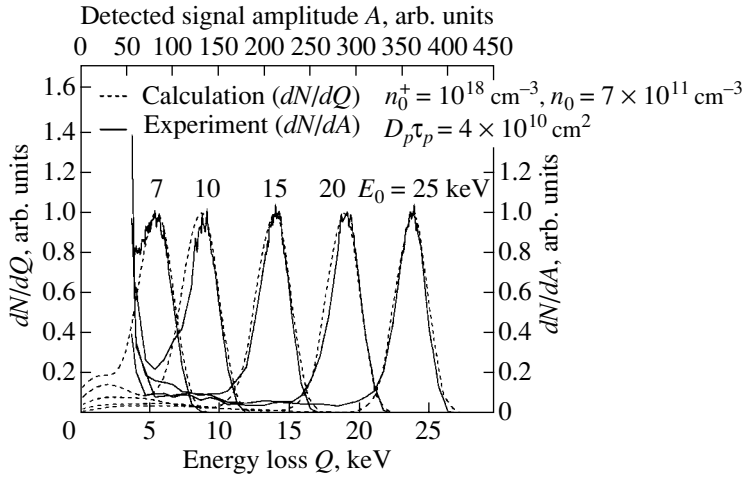


Fig. 5. The comparison of experimental signal spectrum of a detector irradiated with monoenergetic electrons with energies in the range from 7 to 25 keV with the results of simulation with the charge-carrier transport taken into account.

In Fig. 4, we show the dependence of the total current $j(z)$ on the generation depth $z = x_3 - x_0$ at the values of the parameters $p_0 = n_0^+ = 10^{18}$ cm^{-3}, $n_0 = 7 \times 10^{11}$ cm^{-3}, and $\sqrt{D_p\tau_p} = 2 \times 10^{-5}$ cm; the depth distribution of the density of generation of EHPs by electrons with energies of 7 and 20 keV is also shown. It can be seen that, at the characteristic depth of penetration of electrons with energies on the order of 10 keV, the form of $j(z)$ can noticeably affect the detected signal.

The obtained results make it possible to simulate the experimentally observed amplitude distribution of charge pulses. By integrating (with respect to the depth) the distribution of energy lost by an electron along each trajectory with the function $j(z)$ that accounts for the contribution of EHPs to the signal, we can find the amplitude distribution of detected pulses. A variation in the parameters in the expression for $j(z)$ indicates that good agreement between experimental and theoretical dependences is attained if $p_0 = n_0^+ \sim 10^{18}$–10^{20} cm^{-3}, $n_0 = 7 \times 10^{11}$ cm^{-3}, and $\sqrt{D_p\tau_p} = 2 \times 10^{-5}$ cm. A comparison of these dependences for electrons with energies in the range from 7 to 25 keV is shown in Fig. 5. It is noteworthy that the obtained values of the charge-carrier concentrations and diffusion lengths are in satisfactory agreement with the design parameters of the semiconductor detector.

CONCLUSION

We used the Monte Carlo method to simulate numerically the depth distribution of generation density for electron-hole pairs (EHPs) when the semiconductor structure is irradiated with monoenergetic electrons with an energy on the order of 10 keV. It is shown that the form of obtained dependences is governed by the spatial distribution of the energy-loss density for a primary electron and is not affected by specific features of the EHP generation.

We obtained an analytical solution to a system of equations that describe the transport of generated EHPs to electrodes in a planar p^+–n–n^+ semiconductor structure with the n-type region depleted completely of charge carriers. It is shown that the losses to recombination in the course of transport profoundly affect the shape of the detected signal.

The results of numerical simulation of the signal for the detector irradiated with monoenergetic electrons with energies in the range from 7 to 25 keV were compared with experimental data. The values of some poorly known characteristics of the semiconductor structure (such as the diffusion lengths of nonequilibrium charge carriers) were used as adjustable parameters. It is shown that one can obtain good agreement between the results of simulation and experimental data if the losses of charge due to recombination are taken into account using the aforementioned analytical expressions.

Thus, we developed a theoretical procedure that makes it possible to calculate the operational characteristics of semiconductor detectors and, consequently, optimize the detector's parameters when designing the promising semiconductor proportional detectors used in analytical methods.

REFERENCES

1. D. C. Koningsberger and R. Prins, *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS and XANES* (Wiley, New York, 1988).
2. *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy*, Ed. by D. Briggs and M. Seah (Wiley, New York, 1983; Mir, Moscow, 1987).
3. M. V. Koval'chuk and V. G. Kohn, Usp. Fiz. Nauk **149**, 69 (1986) [Sov. Phys. Usp. **29**, 426 (1986)].

4. M. Ya. Amus'ya, *The Photoelectric Effect in Atoms* (Nauka, Moscow, 1987) [in Russian].
5. D. P. Woodruff and T. A. Delchar, *Modern Techniques of Surface Science*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1994; Mir, Moscow, 1989).
6. V. I. Nefedov and V. T. Cherepin, *Physical Methods of Studying Solid Surfaces* (Nauka, Moscow, 1983) [in Russian].
7. http://www.ioffe.ru/XIEES/.
8. Yu. K. Akimov, O. V. Ignat'ev, and A. I. Kalinin, *Semiconductor Detectors in Experimental Physics*, Ed. by Yu. K. Akimov (Énergoatomizdat, Moscow, 1989) [in Russian].
9. *Proceedings of the 9th European Symposium on Semiconductor Detectors, Schloss Elmau, 2002;* http://www.hll.mpg.de/elmau/main.html/.
10. I. V. Eremin, S. G. Konnikov, and K. Yu. Pogrebitskiĭ, Pis'ma Zh. Tekh. Fiz. **29** (24), 27 (2003) [Tech. Phys. Lett. **29**, 1024 (2003)].
11. L. A. Bakaleĭnikov, E. Yu. Flegontova, K. Yu. Pogrebitskiĭ, *et al.*, Zh. Tekh. Fiz. **71** (7), 14 (2001) [Tech. Phys. **46**, 796 (2001)].
12. D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963; Mir, Moscow, 1965).
13. Z.-J. Ding, PhD Thesis (Osaka Univ., Osaka, 1990).
14. *Handbook of Optical Constants of Solids II*, Ed. by E. D. Palik (Academic, New York, 1991).
15. E. Yu. Flegontova, L. A. Bakaleĭnikov, K. Yu. Pogrebitskiĭ, *et al.*, Zh. Tekh. Fiz. **70** (12), 6 (2000) [Tech. Phys. **45**, 1518 (2000)].
16. http://www.ioffe.rssi.ru/ES.
17. V. L. Bonch-Bruevich and S. G. Kalashnikov, *Physics of Semiconductors* (Nauka, Moscow, 1977) [in Russian].
18. G. E. Pikus, *Fundamental Theory of Semiconductor Devices* (Nauka, Moscow, 1965) [in Russian].
19. B. I. Reznikov and G. V. Tsarenkov, Fiz. Tekh. Poluprovodn. (Leningrad) **25**, 1922 (1991) [Sov. Phys. Semicond. **25**, 1158 (1991)].

Translated by A. Spitsyn

Technical Physics, Vol. 49, No. 9, 2004, pp. 1191–1195. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 86–90.

OPTICS, QUANTUM ELECTRONICS

Interference Nanolithography with a UV Laser

V. I. Bredikhin*, V. N. Burenina*, Yu. K. Verevkin*, A. V. Kirsanov*, V. N. Petryakov*, N. V. Vostokov, V. F. Dryakhlushin**, and A. Yu. Klimov****

** Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603950 Russia*

e-mail: verevkin@appl.sci-nnov.ru

*** Institute of Physics of Microstructures, Russian Academy of Sciences, Nizhni Novgorod, 603950 Russia*

Received July 21, 2003; in final form, January 6, 2004

Abstract—The sensitivity and resolution of a photoresist composed of a two-layer (polymer–metallic indium) film are measured. 2D masks used to create nanodimensional metallic and insulating islands on a silicon substrate are prepared by direct laser action. Conditions are found for preparing submicron periodic structures on TiO_2 films that are applied on a glass substrate by the sol–gel technology. Optical properties of these arrays are measured, and it is shown that they can be used for exciting plane electromagnetic waves. *© 2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Interference lithography as a tool for creating large periodic nanometer arrays with different dimension and symmetry has recently found increasing favor [1–6]. Basically, using radiation from an F_2 laser and achromatic interference optics [7, 8], one can localize the laser energy within an ensemble of 40-nm areas with a period of 80 nm. A simple scheme incorporates appropriate phase structures, which can be prepared with longer wavelength (e.g., Xe–Cl) lasers. When the laser radiation diffracts on the phase structure, the period of variation of the intensity in the standing wave can be made twice as short as the period of the phase structure.

Localization of laser radiation in such tiny areas seems promising for creating quantum-size objects, locally enhancing impurity diffusion, generating defects (which is sometimes necessary in the microelectronic technology [9, 10]), etc. The outlook for interference lithography is discussed in [11]. The feasibility of producing germanium quantum dots on a nonplanar silicon substrate that is patterned by holographic lithography combined with reactive ion etching was demonstrated in [12].

Pulsed interference lithography offers wide potentialities for producing large ensembles of subwavelength holes in metal films for studying the effect of resonant plasma oscillations on the optical parameters of the films [13]. Also, the technique discussed can be used for production of nanometer SiO_2 islands in the SOI technology, for local gettering of impurities, etc. The optical properties of the TiO_2 films grown by the sol–gel technology (especially of those doped by various impurities [18–21]) are of interest for use in fiber-optic and planar devices (in particular, in Bragg gratings [14–17]).

EXPERIMENTAL SETUP AND EXPERIMENTAL RESULTS

In this work, we study direct modification of the indium films by four-beam interference, as well as of the TiO_2 films obtained by the sol–gel technology (interference of two UV beams).

The source of UV radiation was a Xe–Cl laser with the following characteristics: the radiation wavelength is 308 nm; the coherence length, more than 30 cm; the pulse width, 7 ns; and the peak energy, up to 100 mJ. The divergence was close to the diffraction limit. The high-power nanosecond UV pulses used in the experiment loosen the requirements for the mechanical stability of the optical system and allow for direct modification of various materials on the nanometer scale [22].

Indium films are of interest, since they can be used as a basis for a fairly sensitive photoresist [22] for which the threshold energy of exposure is virtually independent of the radiation spectrum. Its essential disadvantage is the adverse effect of heat diffusion; hence, pulses shorter than 10 ns should be applied to achieve submicron resolution [23].

The samples used in the experiments consisted of a silicon substrate coated by silicon dioxide (the thickness $h \approx 80$ nm), a polymer ($h \approx 50$ nm), and an indium film ($h \approx 10$ nm). With an indium film 10 nm thick, through holes were obtained at a mean energy density of 20 mJ/cm^2 (Fig. 1). Figure 1 shows the AFM image of a part of the indium film and the section of the hole marked by the bar. The hole has a steep slope and a flat bottom. This is associated with the presence of the polymer underlayer, which features a higher ablation temperature and a lower optical absorption. From these measurements, one can estimate the resolution of the two-layer resist, which turns out to be 100 nm or higher.

1063-7842/04/4909-1191$26.00 © 2004 MAIK "Nauka/Interperiodica"

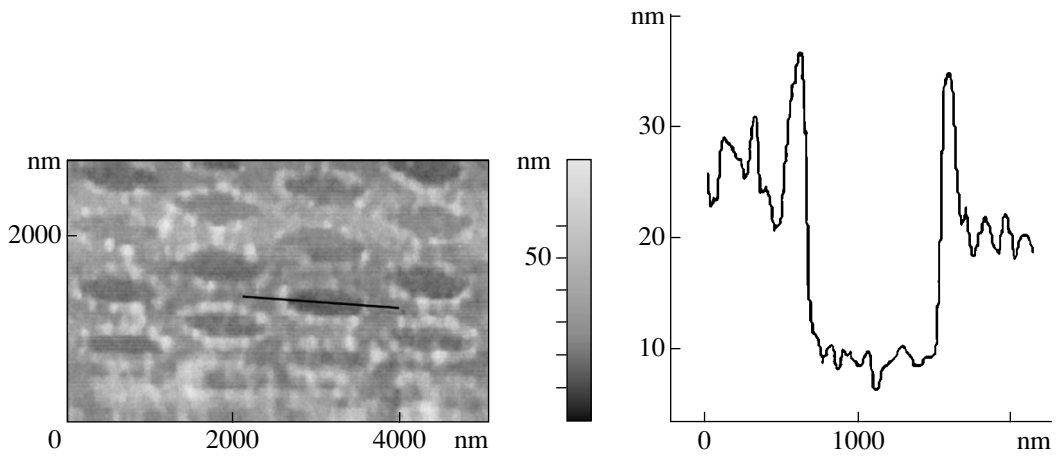


Fig. 1. AFM image of the metallic indium film after irradiation by four coherent beams of the Xe–Cl laser radiation.

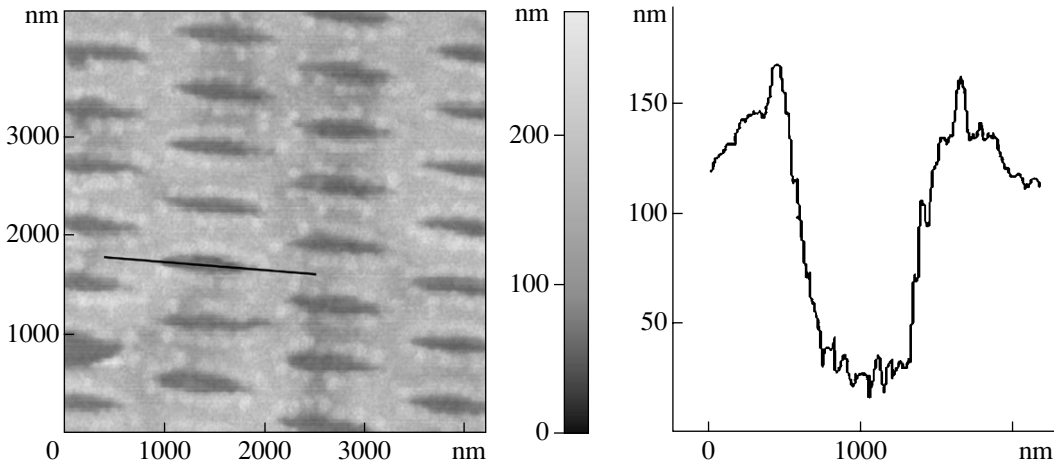


Fig. 2. AFM image of the polymer film after indium mask etching.

Then, the pattern was transferred on the surface of the sample by etching the polymer layer in an oxygen plasma with the unexposed areas remaining intact. This process is well controlled: the related AFM image and the section of the hole are shown in Fig. 2. It is seen that the slope of the hole has become less steep but the hole itself has become much deeper. The decrease in the steepness in this case is probably related to the shape of the AFM tip.

The mask thus formed has a high height-to-width ratio and makes it possible to form metallic and insulating nanofeatures on the silicon surface or apply dry etching with subsequent lift-off lithography. A large array of nanometer metallic contacts on the silicon dioxide surface is shown in Fig. 3. The result of plasma-chemical mask etching of the silicon dioxide layer (an array of nanoholes on the silicon surface) is shown in Fig. 4.

Our next goal was to vary the properties of the TiO_2 films obtained by the sol–gel technology. TiO_2 films are viewed as a promising material for various planar waveguides. The radiation from a Xe–Cl laser and a TiO_2 film interact with a high efficiency, because 308-nm radiation falls into the interband absorption range.

The TiO_2 films were prepared by the chemical method from a 5% solution of $Ti(OC_4H_9)_4$ [24]. The solution was poured over a rotary glass substrate (the rate of rotation was ≈400 rpm). The films were firmly attached to the surface by quenching in an air thermostat at 150°C for a day. To reach a desired thickness of the film, it was applied in several steps, each subsequent layer being heated. The final thickness was varied from 100 to 500 nm. Patterns with periods of 560, 420, and 330 nm were prepared by means of two-beam interference. Examples of such patterns are depicted in Figs. 5 and 6. The threshold energy density of structure modification was ≈30 mJ/cm^2. This allowed us to pattern an area of ≈1 cm^2 per laser shot. It should be noted that if the threshold energy density was exceeded, the patterns had a period twice shorter than the period of

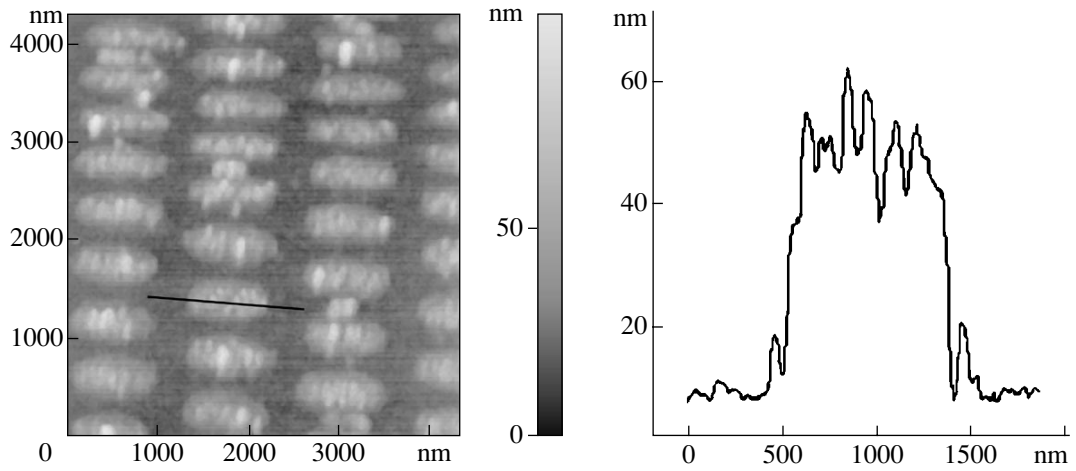


Fig. 3. AFM image of the metal islands on the silicon dioxide surface.

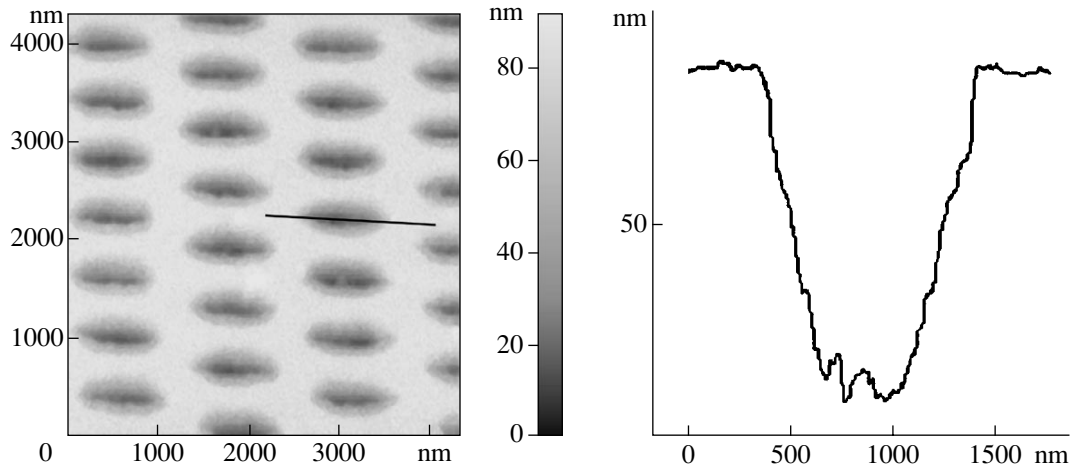


Fig. 4. AFM image of the structure after etching of the silicon dioxide layer to form nanoholes on the sample surface.

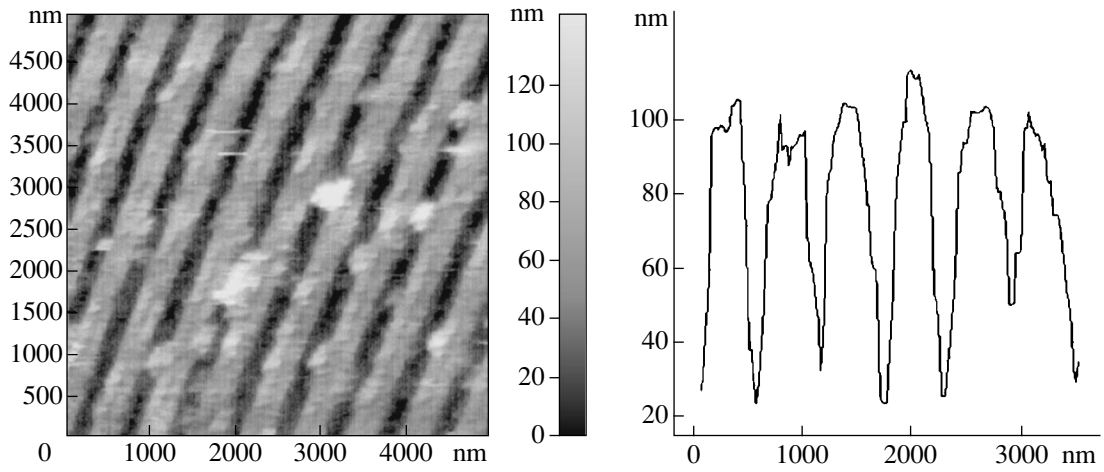


Fig. 5. Patterned TiO_2 film after the action of two coherent UV beams. The period of the standing wave in the plane of the sample is ≈560 nm; the energy density of exposure, ≈30 mJ/cm^2.

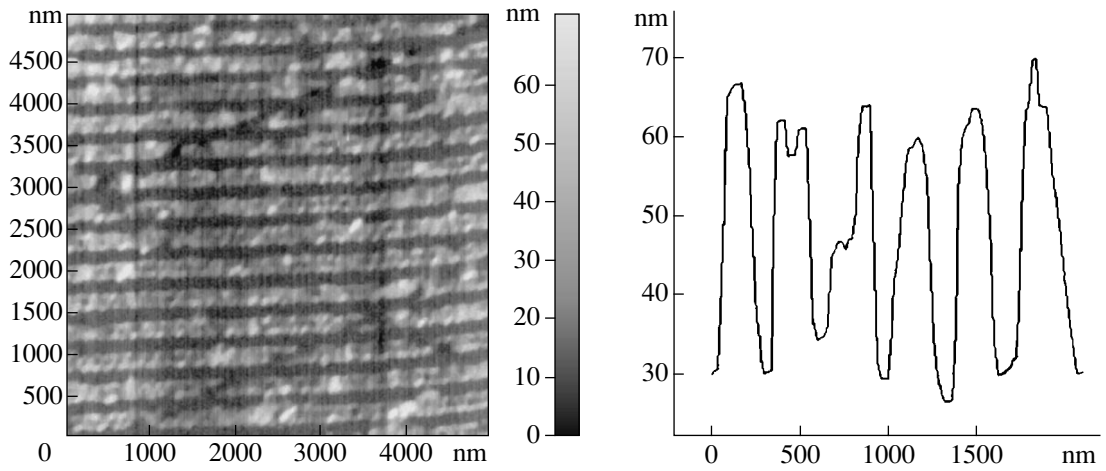


Fig. 6. Patterned of the TiO_2 film after the action of two coherent UV beams. The period of the standing wave in the plane of the sample is ≈330 nm; the energy density of exposure, ≈30 mJ/cm^2.

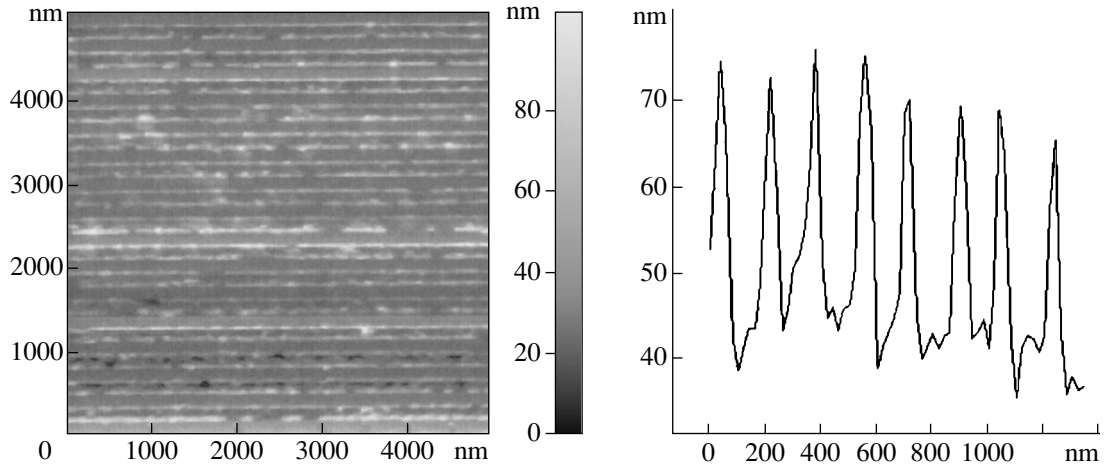


Fig. 7. AFM image of the TiO_2 film surface. The period of the standing wave is ≈330 nm; the energy density of exposure, ≈60 mJ/cm^2.

the standing wave of the radiation. This results from the fact that each of the interference maxima produces two hillocks on the surface in this case, since the melt slides down in opposite directions. Figure 7 exemplifies such a situation. Here, one can see the structure that is composed of two periodic gratings embedded in each other and arranged nearly symmetrically relative to each other.

Using a He–Ne laser, we performed tentative experiments on excitation and reradiation of plane waves through the periodic structures prepared. The samples (gratings) were mounted on a goniometer, and angles of incidence of the radiation at which diffraction waves of different orders arise in the air and in the medium were measured. Note that, if the diffracted waves in the air are absent, counterpropagating or copropagating waves (depending on the angle of incidence and the grating period) may be excited in the medium. It is also easy to create a situation where a surface wave is partially reradiated into the air as a result of diffraction by the periodic grating in the TiO_2. Conditions for increasing the efficiency of radiation transfer from the air to the film will be reported in subsequent publications.

CONCLUSIONS

Pulsed interference lithography equipment that is provided with a variety of optical devices and probe microscopes to characterize films and their surfaces is designed. Varying the number of interfering beams, their polarization, and angles between them, we created structures with different periods and symmetries. The polymer was etched through the 2D indium mask obtained by direct laser action, and then metallic and insulating nanofeatures were formed on the silicon substrate. Submicron-period gratings were made on the

TiO_2 film. Such equipment is promising for research in the field of laser nanotechnology.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 02-02-17639).

REFERENCES

1. O. Beyer, I. Nee, F. Havermeyer, and K. Buse, Appl. Opt. **42**, 30 (2003).
2. B. J. Egglet, IEEE J. Sel. Top. Quantum Electron. **7**, 409 (2001).
3. L. Z. Cai and X. L. Yang, Opt. Laser Technol. **34**, 671 (2002).
4. K. Kintaka, J. Nishii, and N. Tohge, Appl. Opt. **39**, 489 (2000).
5. A. Fernandez, P. J. Bedrossian, S. L. Baker, *et al.*, IEEE Trans. Magn. **32**, 4472 (1996).
6. V. I. Bredikhin, Yu. K. Verevkin, É. Ya. Daume, *et al.*, Kvantovaya Élektron. (Moscow) **30**, 333 (2000).
7. T. Itani and W. Wacamiya, Microelectron. Eng. **61–62**, 49 (2002).
8. M. Switkes, T. M. Blomstein, and M. Rothschild, Appl. Phys. Lett. **77**, 3149 (2000).
9. S. T. Šišianu, T. S. Šišianu, and S. K. Railean, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 611 (2002) [Semiconductors **36**, 581 (2002)].
10. S. V. Vintsent, A. V. Zaĭtsev, and G. S. Plotnikov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **37**, 134 (2003) [Semiconductors **37**, 124 (2003)].
11. H. I. Smith, Physica E **11**, 104 (2001).
12. Z. Zhoug, A. Halilovic, M. Muhlberger, *et al.*, Appl. Phys. Lett. **82**, 445 (2003).
13. T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, *et al.*, Nature **391**, 667 (1998).
14. S. Pelissier, D. Blanc, and M. P. Andrews, Appl. Opt. **38**, 6744 (1999).
15. H. J. Jiang, X. C. Yuan, Y. Zhou, *et al.*, Opt. Commun. **185**, 19 (2000).
16. A. Shishido, I. B. Diviliausky, I. C. Khoo, *et al.*, Appl. Phys. Lett. **79**, 3332 (2001).
17. W. Que and C. H. Kam, Opt. Eng. **41**, 1733 (2002).
18. A. Selvarajian and T. Srinivas, IEEE J. Quantum Electron. **37**, 1117 (2001).
19. A. Howard, D. N. S. Clark, C. E. J. Mitchell, *et al.*, Surf. Sci. **518**, 210 (2002).
20. C. Alcober, F. Alvares, A. A. Bilmes, *et al.*, J. Mater. Sci. Lett. **21**, 501 (2002).
21. S. Komuro, T. Katsumata, and H. Kokai, Appl. Phys. Lett. **81**, 4733 (2002).
22. Yu. K. Verevkin, N. V. Vostokov, and V. F. Dryakhlushin, in *Proceedings of the Conference "Basic Problems of Optics with School for Students, Young Scientists, and Engineers Optics-2000," St. Petersburg, 2000*, pp. 92–93.
23. Yu. K. Verevkin and É. Ya. Daume, Opt. Spektrosk. **85**, 260 (1998) [Opt. Spectrosc. **85**, 239 (1998)].
24. N. V. Suĭkovskaya, *Chemical Methods of Thin Transparent Film Production* (Khimiya, Moscow, 1971) [in Russian].

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 9, 2004, pp. 1196–1201. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 91–96.

RADIOPHYSICS

Attenuation in Waveguide

I. A. Kotel'nikov

Budker Institute of Nuclear Physics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 11, Novosibirsk, 630090 Russia

e-mail: I.A.Kotelnikov@inp.nsk.ru

Received March 11, 2003; in final form, December 29, 2003

Abstract—It is demonstrated that strong absorption of a TM wave upon grazing incidence on a metal surface is analogous to total transmission of a TM wave incident on a dielectric surface at the Brewster angle. Vanishing reflectance of the metal surface leads to increasing attenuation of oscillations propagating in a waveguide. The ratio κ/k of the transverse wave number κ to the limiting wave number k is on the order of the surface impedance ζ. The attenuation coefficient of H and E waves in a circular waveguide is calculated for an arbitrary relation between κ/k and ζ. It is demonstrated that for $\kappa/k \ll \zeta$, the attenuation coefficient is less than that predicted by the theory based on successive approximations (see, for example, §90 in [1]). © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

A conventional approach to the calculation of attenuation of waves in a waveguide with conducting walls employs the method of successive approximations based on the smallness of attenuation of waveguide modes (see, for example, [1]). In this work, by the example of a circular waveguide, we demonstrate that such an approach yields an incorrect result provided that the wavelength is much less than the radius of the cross section and that analysis involves modes with a small transverse wave number. It is demonstrated that such modes can exhibit a field structure that fundamentally differs from the field structure in a waveguide with perfectly conducting walls even in the case of low surface impedance that serves as a small parameter in the method of successive approximations. This phenomenon is closely related to the effect of total transmission of an electromagnetic wave incident on an interface between two media at the Brewster angle (see Section 1). In Section 2, we discuss the attenuation of waves in a circular waveguide. The concluding remarks show that for the major part of applications, the successive approximation procedure yields correct results. We also present examples showing when the results of exact calculations need to be used.

1. REFRACTION AT METAL INTERFACE

Consider the reflection of a plane electromagnetic wave from a metal surface. The amplitudes E_1 and E_2 of refracted and reflected waves can be represented in terms of the incident wave amplitude E_0 using the solution to a similar problem for an insulator and substituting the expression

$$\varepsilon = \frac{4\pi i\sigma}{\omega}, \tag{1}$$

where σ is the conductivity of the metal.

For a TE wave (Fig. 1a), the Fresnel formulas (§86 in [1]) yield the following expressions:

$$\frac{E_1}{E_0} = \frac{\cos\theta - \sqrt{\varepsilon}\cos\theta_2}{\cos\theta + \sqrt{\varepsilon}\cos\theta_2}, \tag{2a}$$

$$\frac{E_2}{E_0} = \frac{2\cos\theta}{\cos\theta + \sqrt{\varepsilon}\cos\theta_2}. \tag{2b}$$

The corresponding relationships for a TM wave are written as

$$\frac{E_1}{E_0} = \frac{\sqrt{\varepsilon}\cos\theta - \cos\theta_2}{\sqrt{\varepsilon}\cos\theta + \cos\theta_2}, \tag{3a}$$

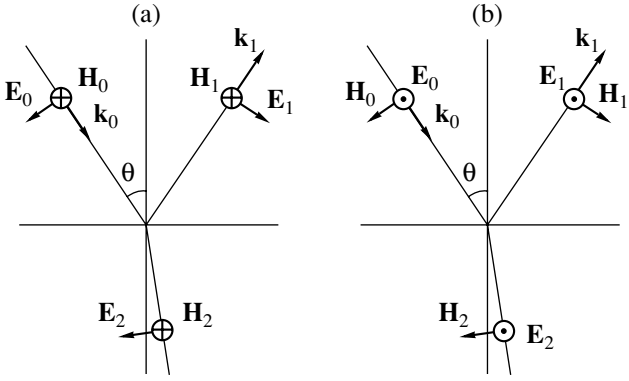


Fig. 1. Reflection of (a) TM and (b) TE waves from plane surface.

1063-7842/04/4909-1196$26.00 © 2004 MAIK "Nauka/Interperiodica"

Parameters of some metals

	σ_1, s^{-1}	σ_0, s^{-1}	ω_p, s^{-1}	τ, s	l_0, cm	δ_0, cm	$\vartheta_{B,\,FIR}$, arcmin	ϑ'_B	$\vartheta_{B,\,NIR}$, deg
Ag	6.1×10^{17}	5.4×10^{17}	1.4×10^{16}	3.6×10^{-14}	5.0×10^{-6}	1.4×10^{-7}	1.8	45	10
Al	3.6×10^{17}	2.1×10^{18}	1.8×10^{16}	8.0×10^{-14}	1.3×10^{-5}	6.1×10^{-8}	0.9	33	7.5
Au	4.4×10^{17}	3.6×10^{17}	1.4×10^{16}	2.4×10^{-14}	3.4×10^{-6}	1.7×10^{-7}	2.2	45	10
Cu	5.8×10^{17}	1.6×10^{17}	1.2×10^{16}	1.3×10^{-14}	1.7×10^{-6}	2.7×10^{-7}	3.4	52	11
Fe	1.0×10^{17}	2.1×10^{16}	5.4×10^{-15}	9.2×10^{-15}	6.9×10^{-7}	1.1×10^{-6}	9.1	120	29
Ni	1.5×10^{17}	6.2×10^{16}	7.4×10^{15}	1.4×10^{-14}	1.3×10^{-6}	5.6×10^{-7}	5.4	86	20
Ti	2.1×10^{16}	8.6×10^{15}	4.5×10^{15}	5.3×10^{-15}	3.5×10^{-7}	1.9×10^{-6}	14	160	38

$$\frac{E_2}{E_0} = \frac{2\cos\theta}{\sqrt{\varepsilon}\cos\theta + \cos\theta_2}. \quad (3b)$$

Here, θ is the angle of incidence and $\cos\theta_2 = \sqrt{1 - \sin^2\theta/\varepsilon}$.

For metals with high conductivity, $|\varepsilon| \gg 1$. Therefore, with a high accuracy, we have $\cos\theta_2 = 1$.

It is known that at metal surfaces, the Leontovich boundary condition (§87 in [1]) must be satisfied for the tangential components of electric and magnetic fields:

$$\mathbf{E}_\tau = \zeta[\mathbf{n}, \mathbf{H}_\tau]. \quad (4)$$

Here, ζ is the surface impedance and $\mathbf{n}$ is the outward normal to metal surface.

With allowance for the fact that $|\mathbf{E}| = |\mathbf{H}|$ outside the metal and using formulas (2) and (3), it can easily be demonstrated that the surface impedance equals $1/(\cos\theta_2\sqrt{\varepsilon})$ $(\cos\theta_2/\sqrt{\varepsilon})$ for a TE (TM) wave. Thus, under the assumption that $\cos\theta_2 = 1$, we obtain

$$\zeta = 1/\sqrt{\varepsilon} = (1 - i)\sqrt{\omega/8\pi\sigma} \quad (5)$$

regardless of the polarization of the electromagnetic field and the angle of incidence. This expression is valid for the normal skin effect, when the thickness of skin layer

$$\delta = \sqrt{c/2\pi\sigma\omega} \quad (6)$$

is much less than the free path of electrons:

$$|l(\omega)| \ll |\delta(\omega)|. \quad (7)$$

Since

$$l(\omega) = l_0/(1 - i\omega\tau), \quad \sigma(\omega) = \sigma_0/(1 - i\omega\tau), \quad (8)$$

we can represent expression (7) as [2]

$$l_0/\delta_0 \ll (1 + \omega^2\tau^2)^{3/4}/(\omega\tau)^{1/2}, \quad (9)$$

where $\delta_0 = c/\omega_p$ and $\omega_p = (4\pi\sigma_0/\tau)^{1/2}$ is the plasma frequency of conduction electrons.

Table shows parameters of several metals. The values of parameters σ_0, τ, ω_p, l_0, and δ_0 are calculated using the results of optical measurements of the reflection of CO_2 laser radiation with a frequency of f = 28.3 THz (λ = 10.6 μm) from metal surfaces [3]. Being applied in a wide (from quasi-stationary fields to visible radiation) frequency range, formulas (8) yield only rough approximation. For example, the values of σ_0 calculated for a few metals substantially differ from the values of static conductivity σ_1 at a temperature of 0°C from [4] (see table). This difference is quite natural, since $l_0/\delta_0 \gg 1$ and the low-frequency range, where the conductivity is close to static, is separated from the IR range by the region of the anomalous skin effect, where formulas (5), (6), and (8) cease to be valid (§86 in [5]).

The real part ζ' of the surface impedance $\zeta = \zeta' + i\zeta''$ is positive (§87 in [1]). In the low-frequency range of the normal skin effect, $\zeta' = -\zeta'' = \sqrt{\omega/8\pi\sigma_1}$, since the static conductivity σ_1 is a real quantity. In the high-frequency range of the normal skin effect, starting from the submillimeter range, where $\omega\tau > 1$, the surface impedance is almost purely imaginary, $\zeta \approx (1 - i\omega\tau)/(\omega\tau)$. In the range of the anomalous skin effect, lying in between the low-frequency and high-frequency ranges of the normal skin effect, $\zeta' \sim \zeta'' \propto \omega^{2/3}$.

It follows from the above facts that the Leontovich boundary condition (4) is satisfied at $|\zeta| \ll 1$. For the subsequent analysis, it is important that this condition can be used at any angles of incidence of radiation on the conducting surface (including the angles corresponding to glazing incidence). We need the well-known facts presented above to draw this conclusion. Inequality $|\zeta| \ll 1$ is satisfied if $\omega \ll \omega_p$. It follows from the data presented in table that the corresponding frequencies can reach the visible range.

We can find the reflection coefficient $R = |E_1/E_0|^2$ for metal surface using the Fresnel formulas (2a) and (3a) assuming that $\cos\theta_2 = 1$ or directly from the Leontovich boundary condition (4) representing E_τ and H_τ in terms

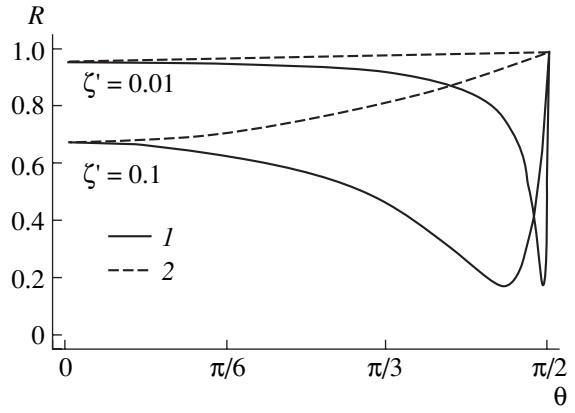


Fig. 2. Plots of reflection coefficient vs. angle of incidence for (*1*) TM and (*2*) TE waves in the case of real conductivity.

of E_0 and E_1:

$$R_{\mathrm{TE}} = \left|\frac{\zeta\cos\theta - 1}{\zeta\cos\theta + 1}\right|^2, \quad (10a)$$

$$R_{\mathrm{TM}} = \left|\frac{\cos\theta - \zeta}{\cos\theta + \zeta}\right|^2. \quad (10b)$$

The reflection coefficient for a TE wave is close to unity at all angles of incidence (Fig. 2), whereas the reflection coefficient of a TM wave exhibits a minimum at $\theta_B = \pi/2 - |\zeta|$:

$$R_{\min} = \frac{|\zeta| - \zeta'}{|\zeta| + \zeta'}.$$

At frequencies for which we can neglect the dispersion of conductivity, $R_{\min} = (\sqrt{2} - 1)/(\sqrt{2} + 1) \approx 0.17$. When the imaginary part ζ'' of the surface impedance decreases, $R_{\min}$ tends to zero and the corresponding angle of incidence tends to the Brewster angle (§86 in [1]). A TM wave incident on a dielectric (metal) surface at the Brewster angle is completely transmitted (absorbed).

For real metals, the difference (equal to $|\zeta|$) between the Brewster angle θ_B and the right angle is very small. Therefore, below, we measure the angle of incidence from the plane of interface and introduce angle $\vartheta = 90° - \theta$. At frequency $f = \omega/2\pi = 1$ MHz, the Brewster angles are about $0.2''$ and no greater than $1''$ for copper and steel, respectively. The surface impedance increases with frequency. However, the impedance is low even in the IR range. For frequency $f = 0.3$ THz ($\lambda = 1$ mm) corresponding to the boundary of far IR range, the Brewster angle at metal surfaces ranges from $1'$ to $10'$. In the near IR range ($f = 375$ THz and $\lambda = 0.8$ μm), the Brewster angle can be as high as a few tens of degrees. The calculated values of $\vartheta_{B,\,FIR}$ and $\vartheta_{B,\,NIR}$ expressed in angular minutes and degrees, respectively, are presented in the third from last and the last columns of table. The second from last column shows the values of the Brewster angle for the CO_2 laser frequency (28.3 THz).

2. ATTENUATION IN CIRCULAR WAVEGUIDE

A deep valley in the reflectance of a metal surface leads to a strong absorption of oscillations in a waveguide at a certain relation between the wavelength, the linear sizes of the waveguide cross section, and the surface impedance of walls. Consider a waveguide with circular (radius a) cross section. In this waveguide, the fields are given by (§91 in [1])

$$\begin{aligned}
\mathbf{E} &= \frac{ik}{\kappa^2}\nabla_\perp E_z + E_z\hat{\mathbf{z}} + \frac{i\omega}{c\kappa^2}[\nabla_\perp H_z, \hat{\mathbf{z}}],\\
\mathbf{H} &= \frac{ik}{\kappa^2}\nabla_\perp H_z + H_z\hat{\mathbf{z}} - \frac{i\omega}{c\kappa^2}[\nabla_\perp E_z, \hat{\mathbf{z}}],\\
E_z &= E_0 J_m(\kappa r)e^{ikz + im\varphi - i\omega l},\\
H_z &= H_0 J_m(\kappa r)e^{ikz + im\varphi - i\omega l},
\end{aligned} \quad (11)$$

where $\kappa = (\omega^2/c^2 - k^2)^{1/2}$ and J_m is the Bessel function.

Using boundary condition (4), we find the dispersion relation [5, 6]

$$\left[\zeta J_m + \frac{i\omega}{c\kappa}J'_m\right]\left[J_m + \zeta\frac{i\omega}{c\kappa}J'_m\right] = -\zeta\left[\frac{ik}{\kappa}\frac{im}{\kappa a}J_m\right]^2, \quad (12)$$

where functions J_m and J'_m are calculated for argument κa.

At the given frequency ω, this equation yields the longitudinal wave number k.

For the ideal conduction of waveguide walls ($\zeta = 0$), Eq. (12) is split into two equations making it possible to independently determine amplitudes E_0 and H_0:

$$J_m(\kappa a) = 0 \quad (E_0 \neq 0, H_0 = 0), \quad (13a)$$

$$J'_m(\kappa a) = 0 \quad (E_0 = 0, H_0 \neq 0). \quad (13b)$$

The former equation corresponds to an E wave with $H_z \equiv 0$. The latter equation corresponds to an H wave with $E_z \equiv 0$. Note that for $m = 1$, the first three roots $\kappa_E a$ ($\kappa_H a$) of Eq. (13a) (Eq. (13b)) are 3.83171, 7.01559, and 10.1735 (1.84118, 5.33144, and 8.53632).

At a finite impedance, any waveguide mode exhibits six components of electromagnetic field (except for the case $m = 0$; see below). However, at a low impedance, one polarization remains dominating. For $|\zeta| \ll \kappa/k$, we obtain from Eq. (12) the following expression for an E wave:

$$J_m(\kappa a) = -\zeta\frac{i\omega}{c\kappa}J'_m(\kappa a).$$

Using successive approximations, we solve this equation and find correction $\delta\kappa$ related to the finiteness of impedance:

$$\delta\kappa = -\frac{i\omega}{c\kappa_E}\frac{\zeta}{a}.$$

Here, κ_E is the root of Eq. (13a). At a fixed frequency ω, $\kappa\delta\kappa + k\delta k = 0$. Then, knowing the calculated variation $\delta\kappa$ in the transverse wave vector, we find the attenuation coefficient of an E wave:

$$\mathrm{Im}k = \frac{\omega}{kc}\frac{\zeta'}{a}, \qquad (14a)$$

where Im denotes imaginary part.

A similar approximate equation for an H wave is written as

$$J'_m = \zeta\frac{i\kappa c}{\omega}J_m + \zeta\frac{ikc}{\omega}\frac{km^2}{\kappa^2 a^2}J_m.$$

Using this expression, we find that

$$\mathrm{Im}k = \frac{kc}{\omega}\frac{m^2k^2 + \kappa_H^4 a^2}{(\kappa_H^2 a^2 - m^2)k^2}\frac{\zeta'}{a}, \qquad (14b)$$

where κ_H is the root of Eq. (13b) such that inequality $\kappa_H a > |m|$ is always satisfied.

Note a distinctive feature of an H wave with $m = 0$: at $\kappa/k \ll 1$, its attenuation coefficient estimated as $(\kappa/k)^2(\zeta'/a)$ is less than the attenuation coefficients for other waves by a factor of $(\kappa/k)^2$.

It follows from the definition of attenuation coefficient that the energy flux of the wave decreases proportionally to $\exp(-2\mathrm{Im}kz)$. Formulas (14) coincide with the classical formulas from textbooks on electrodynamics (see, for example Problem 2 from §91 in [1]), where the dissipation in waveguide walls serves as a small perturbation. However, these formulas yield incorrect result if $\kappa/k \lesssim |\zeta| = \kappa_B$.

To demonstrate this, we consider the case $m = 0$. For an arbitrary relation between ζ and κ/k, Eq. (12) is split into two independent equations written as

$$J_0 + \zeta\frac{i\omega}{c\kappa}J'_0 = 0 \quad (E_0 \neq 0,\ H_0 = 0), \qquad (15a)$$

$$\zeta J_0 + \frac{i\omega}{c\kappa}J'_0 = 0 \quad (E_0 = 0,\ H_0 \neq 0). \qquad (15b)$$

In the limiting case $\zeta \gg \kappa/k$, we find from these equations that

$$\delta\kappa = -\frac{i\kappa_H}{ka}\frac{1}{\zeta}, \quad \mathrm{Im}k = \frac{\kappa_H^2}{k^2 a|\zeta|^2}\zeta', \qquad (16a)$$

$$\delta\kappa = -\frac{i\kappa_H}{ka}\zeta, \quad \mathrm{Im}k = \frac{\kappa_H^2}{k^2 a}\zeta' \qquad (16b)$$

for E and H waves, respectively. Note that in both cases, in the zero approximation, we find κ_H from one equation $J'_0(\kappa_H a) = 0$. As in the case when $|\zeta| < \kappa/k$, the E wave with $m = 0$ decays faster: the attenuation coefficient given by expression (16a) is greater than that given by expression (16b) by a factor of $1/|\zeta|^2$.

For $m \neq 0$, we leave in expression (12) only the main terms in the limiting case $|\zeta| \gg \kappa/k$:

$$\zeta\left[\frac{i\omega}{c\kappa}J'_m\right]^2 = \zeta\left[\frac{ik}{\kappa}\frac{im}{\kappa a}J_m\right]^2.$$

This equation yields only purely real solutions, since after cancelling out of multiplier ζ, it is independent of the conductivity of walls. The equation can also be split into two independent equations, which can be further simplified taking into account the approximate equality $\omega/c \approx k$ and the inequalities $\kappa/k \ll |\zeta| \ll 1$:

$$\frac{m}{\kappa a}J_m \mp J'_m = J_{m\pm1} = 0 \quad (H_0 = \pm iE_0). \qquad (17)$$

Here, two variants to choose the sign correspond to two equations and the waves with left (upper sign) and right (lower sign) circular polarizations in the plane of the waveguide cross section:

$$E_\tau = \mp\frac{ik}{\kappa}J_{m\pm1}(\kappa\tau)E_0,$$

$$E_\varphi = -\frac{\kappa}{\kappa}J_{m\pm1}(\kappa r)E_0, \qquad (18)$$

$$E_z = J_m(\kappa r)E_0.$$

Note that $\mathbf{H} = \pm i\mathbf{E}$.

Following [6], we name the two solutions (18), corresponding to the upper and lower signs, L and R waves. At the waveguide walls, only the z components of the electric and magnetic fields of these waves differ from zero. Therefore, the Poynting vector $\mathbf{S} = (c/4\pi)[\mathbf{E}, \mathbf{H}]$ equals zero, and the energy dissipation is absent. L and R waves represent a generalization of nondiffracting Bessel beams (exhibiting the absence of time average energy flux from the beam axis) [7] to the case $m \neq 0$. For the case $m = 1$, the first three roots $\kappa_L a$ ($\kappa_R a$) of equation $J_2 = 0$ ($J_0 = 0$) are 5.13562, 8.41724, and 11.6198 (2.40483, 5.52008, and 8.65373).

A correction to quantity $\kappa_{L,R}$ is calculated as

$$\delta\kappa = -\frac{i\kappa_{L,R}}{2ka}\left[\zeta + \frac{1}{\zeta}\right] \pm \frac{\kappa_{L,R}^2}{2k^2 a}\left[\frac{m}{\kappa_{L,R}a} - \frac{\kappa_{L,R}a}{m}\right],$$

and the attenuation coefficient is given by

$$\mathrm{Im}k = \frac{\kappa_{L,R}^2\zeta'}{2k^2 a|\zeta|^2}, \qquad (19)$$

where $\kappa_{L,R}$ can be found from Eq. (17) with regard to the sign.

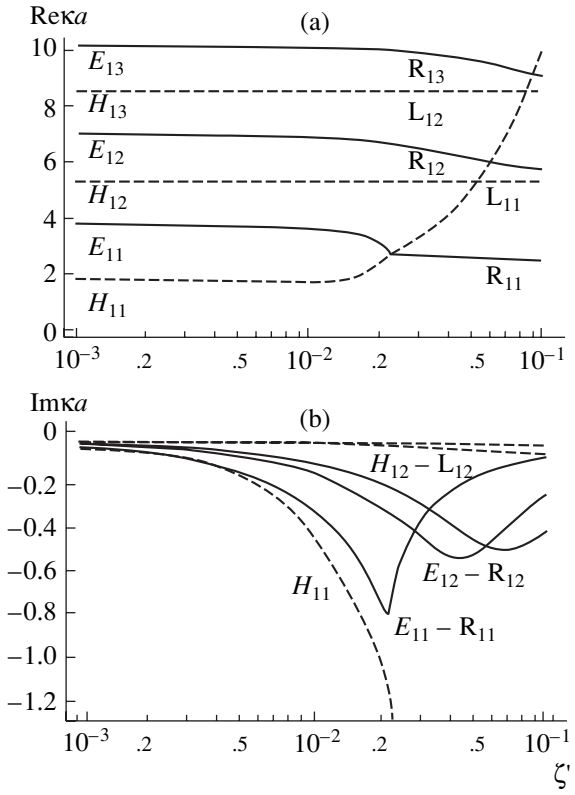


Fig. 3. Plots of (a) real and (b) imaginary parts of κa vs ζ' at $\kappa a = 100$ and a real conductivity. H_{mn}, E_{mn}, L_{mn}, and R_{mn} are waves with azimuthal number m and radial number n.

The maximum attenuation is reached at the common boundary of domains of applicability for formulas (14) and (19) at $|\zeta| \sim \kappa/k$. The corresponding attenuation length $(\mathrm{Im}k)^{-1} \sim a/\zeta'$ can be close to the diffraction spreading length a^2/λ of a wave packet propagating in an infinite space provided that $\zeta' \sim |\zeta|$ (recall that $\zeta' \ll |\zeta|$ in the IR range).

By way of an example of waves with $m = 1$ (Fig. 3), we can trace the mechanism of transformation of H and E waves into L and R waves with the variation in the surface impedance. For simplicity, we plot curves using the equality $\zeta = (1 - i)\zeta'$, which is valid at purely real conductivity coefficient. When ζ' increases from the limiting value $\zeta' \ll \kappa/k$ to $\zeta' \gg \kappa/k$, κ eventually decreases. If we arrange in ascending order, on the one hand, κ_H and κ_E and, on the other hand, κ_R and κ_L, then the relation between the waves at $\zeta' \ll \kappa/k$ and $\zeta' \gg \kappa/k$ can be established using a simple rule: the root from the first sequence is transformed into the nearest smaller root from the second sequence. The smallest (in both sequences) root 1.84118 corresponds to the H_{11} wave. Therefore, this wave vanishes (strongly decays) when the impedance increases. The remaining H waves are transformed into L waves with the radial number n decreased by unity. In particular, H_{12} is transformed into L_{11}, H_{13} is transformed into L_{12}, etc. E waves are transformed into R waves with the same radial number (E_{mn} is transformed into R_{mn}).

CONCLUSIONS

Based on the above analysis, we conclude that the conventional approach to the calculation of attenuation of waves in a waveguide may yield incorrect results, since the finiteness of the resistance of waveguide walls at $|\zeta| \gtrsim \kappa/k$ leads to a structural transformation of waveguide fields. Recall that the conventional approach is based on successive approximations. In the first approximation, we neglect the resistance of walls and find the eigenmodes of a waveguide with perfectly conducting walls assuming that the tangential projection of electric field equals zero ($\mathbf{E}_\tau = 0$) at the surface of the walls. In the next approximation, we calculate E_τ using the Leontovich boundary condition (4) and the value of the tangential projection of magnetic field $\mathbf{H}_\tau$ at the walls found at the first stage. Then, we calculate the energy flux to the walls $\mathbf{S}_n = (c/4\pi)[\mathbf{E}_\tau, \mathbf{H}_\tau]$ and find the attenuation length from the equation for energy balance in the wave.

Under real conditions, the surface impedance ζ is low and condition $|\zeta| \ll \kappa/k$ is always satisfied. Hence, successive approximations yield correct results. Nevertheless, we can indicate two phenomena, in which the above condition can be violated.

Recently, Lotov [8] proposed that microcapillary tubes could be used to prevent diffraction spreading of a laser pulse that serves as a driver in the wake acceleration of charged particles. A laser pulse with a power of tens of terawatts virtually instantaneously ionizes the walls of a microcapillary, transforms any material into a highly conducting substance, and propagates in the capillary as in a waveguide. Using a capillary with appropriate parameters, we can hypothetically form a quasi-Bessel beam and substantially increase the length of spreading.

Another possible application is related to the analysis of the frequency dependence of the longitudinal impedance in cyclic particle accelerators [9, 10]. This dependence exhibits resonance character in the vicinity of frequencies satisfying the condition for time synchronism $\omega = n\omega_0$ (n is integer) involving the frequency ω of synchrotron modes and the rotation frequency $\omega_0 = \beta c/R$ of a particle in the accelerator (βc is the velocity of the particle and R is the radius of its orbit). Synchrotron modes are excited in a toroidal camera by accelerated particles provided that the phase velocity of the mode is less than the velocity of light. The last condition leads to the inequality $\omega > \omega_{\min} \sim \pi c R^{1/2}/a^{3/2}$ [11]. The width of an individual resonance is proportional to the attenuation coefficient $\mathrm{Im}k$ of the resonant mode: $\Delta\omega = \mathrm{Im}kc$. We can estimate the frequency $\omega_{\max}$ corresponding to the transition between the attenuation modes (14) and (19) equating the characteristic propa-

gation angle of synchrotron radiation $\kappa \sim (\omega_0/\omega)^{1/3}$ (for the low-frequency part of the synchrotron radiation spectrum) to the Brewster angle $|\zeta| \sim \omega/\omega_p$. For typical parameters of sources of synchrotron radiation [12], the value $\omega_{min} \sim \omega_p^{3/4}\omega_0^{1/4}$ falls into the IR range and satisfies the condition $\omega_{max} > \omega_{min}$.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon Press, New York, 1984).
2. M. I. Kaganov, G. Ya. Lyubarskiy, and A. G. Mitina, Phys. Rep. **288**, 291 (1997).
3. A. M. Prokhorov, V. I. Konov, I. Ursu, and I. N. Mikhalesku, *Interaction of Laser Radiation with Metals* (Nauka, Moscow, 1988), p. 6 [in Russian].
4. *Tables of Physical Quantities*, Ed. by I. K. Kikoin (Atomizdat, Moscow, 1976) [in Russian].
5. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).
6. I. A. Kotel'nikov, Preprint No. 98-85, IYaF (Budker Institute of Nuclear Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 1989).
7. J. Durnin, J. J. Miceli, and J. H. Eberly, Phys. Rev. Lett. **58**, 1499 (1987).
8. K. V. Lotov, Laser Part. Beams **19**, 219 (2001).
9. King Yuen Ng, Part. Accel. **25**, 153 (1990).
10. R. L. Warnock and P. Morton, Part. Accel. **25**, 113 (1990).
11. G. V. Stupakov and I. A. Kotelnikov, Phys. Rev. ST Accel. Beams **6**, 034401 (2003).
12. V. A. Chernov, V. I. Kondratev, V. N. Korchuganov, *et al.*, Nucl. Instrum. Methods Phys. Res. A **45**, 179 (1998).

Translated by A. Chikishev

Technical Physics, Vol. 49, No. 9, 2004, pp. 1202–1207. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 97–102.

ELECTRON AND ION BEAMS, ACCELERATORS

Characteristics of an Ion Source with a Plasma Cathode and a Multipole Magnetic System for Confining Fast Electrons

N. V. Gavrilov and A. S. Kamenetskikh

Institute of Electrophysics, Russian Academy of Sciences, Ural Division, ul. Komsomol'skaya 34, Yekaterinburg, 620016 Russia

e-mail: gavrilov@iep.uran.ru

Received January 8, 2004

Abstract—Parameters and ion-emission characteristics of the plasma generated in the anode stage of an ion source with a hollow glow-discharge plasma cathode are studied. To decrease the minimum operating gas pressure to 5×10^3 Pa, a multipole magnetic system was installed on the surface of the hollow cathode and the peripheral magnetic field was enhanced in the anode stage of the source. The effect of the gas pressure, the plasma-cathode current, and the voltage between the electrodes of the anode stage on the value of the ion current extracted from the plasma is investigated. It is found that the size of the exit aperture of the hollow cathode substantially affects the efficiency of ion extraction. The potential (1–5 V) and the electron temperature (1–8 eV) of the anode-stage plasma are measured by the probe method. The conditions are determined that ensure the maximum ion-emission current from the plasma at low gas pressures. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The use of a two-stage glow discharge (in the first stage, a plasma cathode is produced and, in the second stage, a plasma with desired ion-emission characteristics is generated) in a gas-ion source not only ensures the long lifetime of the source when operating with chemically active gases, but also makes it possible to decrease the operating gas pressure and the content of metal ions in the ion beam [1]. The use of a grid that separates the cathode from the anode stage of the source and fixes the position and potential of the plasma electron emitter allows one to vary the energy of the electrons injected into the second stage [2]; this offers an additional possibility of controlling the density of the emitting plasma, as well as the mass and charge composition of ions.

In the second stage of the source, various electromagnetic traps can be used that ensure the confinement and efficient energy relaxation of the primary electrons. Moreover, to generate a wide ion beam in such a trap, a spatially uniform plasma should be produced. An important requirement for the confinement system is the possibility of extracting a substantial fraction of the ions produced in the plasma. At present, systems with a multipole magnetic field [3] best meet the above requirements. These systems ensure higher efficiency of ion extraction as compared to systems with the electrostatic confinement of electrons in a hollow cathode [2] and do require special measures for spatially equalizing the plasma density, as in Penning systems [4].

The present paper is devoted to studying the parameters and ion-emission characteristics of the plasma generated in the nonmagnetized second stage of a gas-discharge system with a low-pressure glow-discharge plasma cathode.

EXPERIMENTAL TECHNIQUE

The electrode system of the ion source (Fig. 1) consisted of a hollow cathode, on the axis of which a rod igniting electrode was installed, and an anode chamber. The cathode and anode electrodes were made of stainless steel and had the same diameter of 130 mm. An electrically insulated tungsten wire grid, whose cell

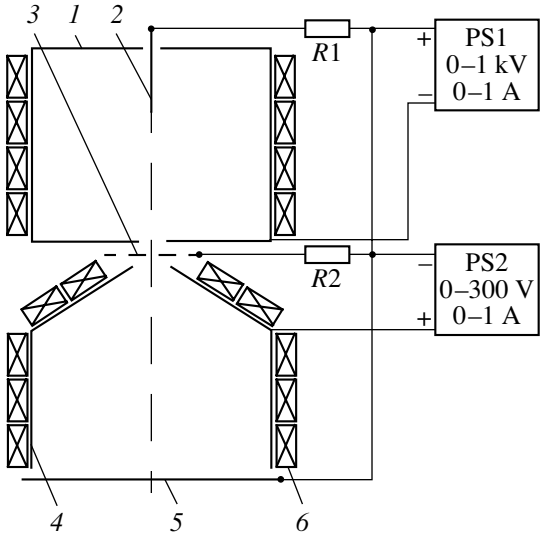


Fig. 1. Electrode system of the ion source: (*1*) hollow cathode, (*2*) igniting electrode, (*3*) grid, (*4*) anode chamber, (*5*) screening plate, and (*6*) magnets.

1063-7842/04/4909-1202$26.00 © 2004 MAIK "Nauka/Interperiodica"

size was varied from 0.2 to 0.6 mm, was placed in front of the exit aperture of the hollow cathode. The aperture diameter was varied from 5 to 20 mm. The ions either were extracted from the plasma through 5-mm holes in the screening electrode of a two-electrode ion-optical system connected electrically to the grid or arrived at a plate placed instead of the screening electrode. The main difference between this electrode system and the system described in [3] is the use of permanent magnets with a large inductance at the pole surfaces (0.15 T) in the second stage, as well as the arrangement of magnets at the surface of the hollow glow-discharge cathode. In experiments, we used $20 \times 10 \times 10$-mm magnets made of the KS37 samarium–cobalt alloy. The multipole magnetic field [5] near the anode surface was generated by 12 longitudinal rows of magnets. The magnets at the cathode were placed in eight rows. The multipole magnetic field at the cathode made it possible to decrease both the working-gas pressure and the operating voltage of the hollow-cathode glow discharge [6]. As a result, it was possible to decrease the working-gas pressure in the vacuum chamber to 5×10^{-3} Pa when operating with the ion source. A stronger magnetic field at the anode improves the confinement of fast electrons [7]; this allows one to achieve efficient ion generation at low pressures.

The cathode cavity was filled with argon. The gas flow rate was varied from 5 to 20 cm^3/min. A hollow-cathode glow discharge was ignited at a voltage of 2.5 kV supplied from power source PS1. The operating voltage of the discharge was 350–650 V at a current of 0.1–1 A. Power source PS2, whose output voltage was varied from 0 to 300 V, ensured the extraction of electrons from the cathode plasma, their acceleration in the cathode sheath formed near the grid, and the extraction of ions from the anode plasma.

Plasma parameters were measured with the help of an emission probe [8], which was a tungsten wire with a diameter of 30 μm and length of 6–8 mm. The wire was welded to the current-carrying conductors passing through channels in a ceramic tube with an outer diameter of 5 mm. In experiments, the probe was placed on the system axis, halfway between the screening electrode and the grid. The probe was heated by a direct current up to a temperature ensuring thermionic electron emission. At an emission current of ~10 mA, the voltage drop across the probe was no higher than 0.5 V. The plasma potential was measured from the position of the inflexion point in the probe characteristic in the regime of low emission currents [9].

To verify the results of measurements of the plasma potential and to determine the energy spectrum of the ions extracted from the plasma, we used a multigrid electrostatic energy analyzer [10]. The analyzer consisted of an ion collector and three sequentially arranged grids with a cell size of 0.1×0.1 mm. The first grid was at the potential of the screening electrode of the ion source, the accelerating voltage was applied to the second grid, and the third grid served as an analyzer. As the potential of the third grid was varied from the screening-electrode potential to a nearly anode potential, the ion retarding curve was measured in the collector circuit. It is well known that the derivative of the retarding curve determines the ion distribution over the longitudinal energy and the position of the maximum of the derivative corresponds to the plasma potential [11]. To improve the accuracy of differentiation and to simplify the measurement procedure, we performed the electronic differentiation of the retarding curve. Along with the dc voltage, a ~1-V modulation voltage was applied to the analyzing grid and the alternating component of the signal at the modulation frequency (2 kHz) was measured in the collector circuit with the help of a selective amplifier. The electronic differentiation was also used to find the inflection point of the emission probe characteristic. For this purpose, the potential applied to the probe was modulated. The dc probe potential at which the amplitude of the first harmonic of the periodic signal in the probe circuit was at maximum was taken as the plasma potential.

The electron temperature was determined from the slope of the linear segment of the cold probe characteristic plotted on a semilogarithmic scale. To decrease the effect of the probe surface contamination on the measurement results, as "cold" probe, we used a heated probe with a temperature lower than that at which electron emission takes place.

RESULTS AND DISCUSSION

An analysis of the processes in a gas-discharge system with a multipole magnetic field [12–14] shows that the ratio I_e/I_i of the current of fast electrons to the current of the ions arriving at the screening plate should be a linear function of the reciprocal of the gas pressure p. The slope of this linear dependence should be inversely proportional to the confinement time of fast electrons, and the intersection point of the extension of this straight line with the ordinate axis determines the ratio of the total electron energy loss in inelastic collisions to the electron energy spent on gas ionization. A general feature of all the experimental results is that an increase in the magnetic field results in a decrease in the slope of the I_e/I_i curve. This testifies to the better confinement of the fast electrons. However, the experimental results have shown that the above dependences measured over a wide range of $1/p$ differ from linear functions.

Figure 2a shows the dependences of the current ratio I_e/I_i measured for two diameters of the exit aperture of the hollow cathode: (*1*–*3*) 20 and (*4*–*6*) 5 mm. The ratio I_e/I_i, which is inversely proportional to the efficiency of ion extraction α, decreases monotonically with increasing pressure; however, after the pressure reaches ~5×10^{-2} Pa, the current ratio sharply increases. The position of the minimum of the I_e/I_i curve depends on the plasma cathode current, the electron energy, and the diameter

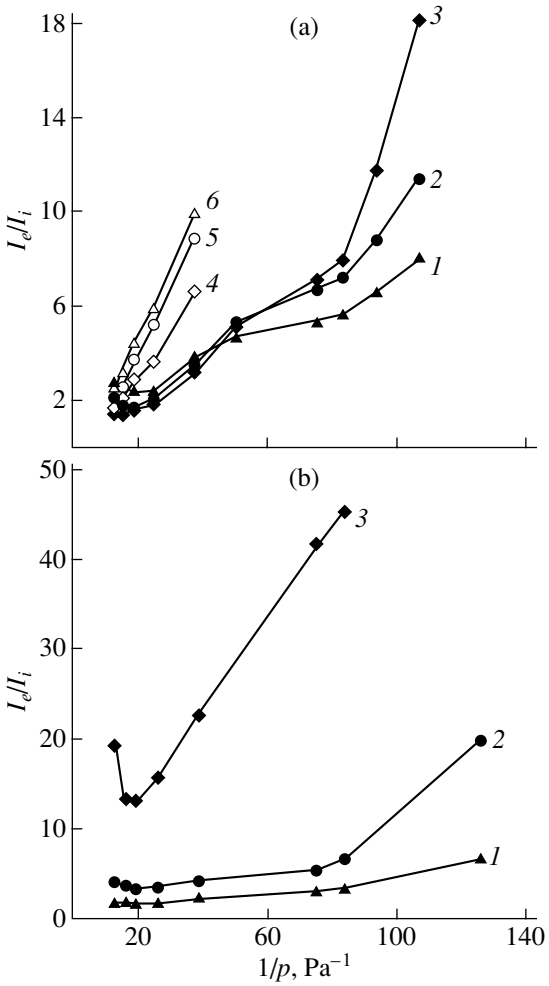


Fig. 2. Ratio between the current of fast electrons I_e and the ion current to the screening plate I_i as a function of the reciprocal of the gas (argon) pressure in the vacuum chamber. (a) The energy of fast electrons is 150 V; the diameter of the exit aperture of the hollow cathode is (*1–3*) 20 and (*4–6*) 5 mm; and the current of fast electrons is I_e = (*1*, *6*) 0.6, (*2*, *5*) 0.4, and (*3*, *4*) 0.2 A. (b) The diameter of the exit aperture of the hollow cathode is 20 mm; the current of fast electrons is I_e = 0.5 A; and the energy of fast electrons is (*1*) 50, (*2*) 100, and (*3*) 200 eV.

of the cathode aperture. The higher the current, the smaller the value of α and the lower the gas pressure at the minimum of the curve. In contrast, in the range of low pressures (curves *1–3*), the efficiency of ion extraction increases with increasing electron current. The minima of curves *4–6* (not shown in Fig. 2a) are observed at higher pressures than for curves *1–3*. As the energy of fast electrons increases (Fig. 2b), the slope of the curves changes and they shift toward smaller values of I_e/I_i. At low electron energies (≈50 eV), the efficiency of ion extraction is low and, in the range of high pressures, the I_e/I_i curve has a sharp minimum.

As the diameter of the exit aperture of the hollow cathode decreases, the efficiency of ion extraction sharply decreases (Fig. 2a). The aperture diameter determines the value of the gas pressure in the cathode cavity at a given gas flow rate. For this reason, when the aperture diameter was increased from 5 to 20 mm, we had to increase (by ~50 V) the voltage between the electrodes of the cathode stage in order to maintain the glow-discharge current and the emission current of the plasma cathode at a constant level. In addition, the ratio of the effective surface area S_a of the anode (which is the grid) to the surface area S_c of the hollow cathode substantially influenced the operating voltage of the hollow-cathode discharge. When the ratio S_a/S_c became higher than $(m_e/M_i)^{1/2}$ (where m_e and M_i are electron and ion masses, respectively), the discharge voltage increased because the loss rate of fast electrons increased in the first stage of the discharge [15]. It should be noted that variations in the voltage between the electrodes of the second stage do not change the operating conditions of the hollow-cathode glow discharge in the first stage of the source. The current–voltage characteristics of this discharge are shown in Fig. 3. When operating at a low gas flow rate and a high glow-discharge voltage, the central region of the grid was heated to fairly high temperatures; this could result in the melting of the stainless-steel grid. The reason for this phenomenon is to be studied. We assume, however, that this occurs because the electron temperature increases or the anode potential drop appears as the cathode sheath length increases and the effective area of the glow-discharge anode decreases. All the results presented below were observed in an electrode system with a 20-mm aperture.

One of the main factors determining the efficiency of the ion source is the plasma potential ϕ_p relative to the anode. When ϕ_p is negative, the ions do not escape to the anode and the initial energy of the fast electrons oscillating in the nonmagnetized plasma decreases. At a positive plasma potential, the ion current at the anode depends on the electron temperature T_e, which (along with the ion temperature) affects the loss area of the plasma particles that return along the magnetic field lines. According to [7], this area is determined by the product of the total length of the magnetic poles by the hybrid gyroradius $(r_e/r_i)^{1/2}$, where r_e and r_i are gyroradii of the plasma electrons and ions, respectively. Figures 4 and 5 show the dependences of the plasma potential and the electron temperature on the discharge current at different gas pressures. It can be seem from these dependences that, over a wide range of the discharge parameters, the plasma potential relative to the anode remains positive and varies within 2–5 V and the electron temperature is 1–8 eV. As the pressure increases, both the plasma potential and the electron temperature decrease. An increase in the current within 0.2–0.8 A results in a slight increase in both the electron temperature and the plasma potential. As the voltage between the electrodes of the second stage increases by 100 V, the electron temperature increases by ~2 eV and the plasma poten-

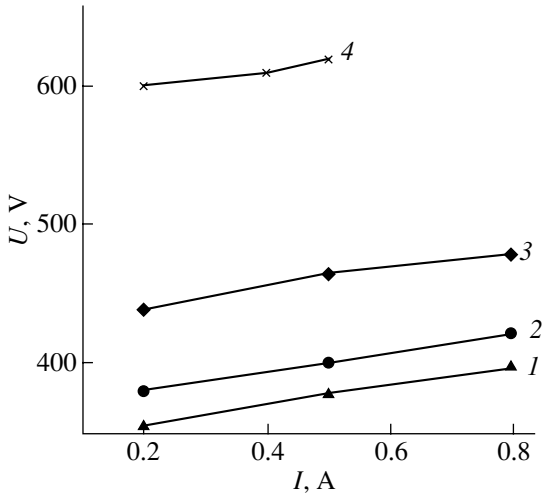


Fig. 3. Current–voltage characteristics of a hollow-cathode glow discharge. The pressure in the vacuum chamber is (*1*) 8.0×10^{-2}, (*2*) 5.3×10^{-2}, (*3*) 2.7×10^{-2}, and (*4*) 0.8×10^{-2} Pa.

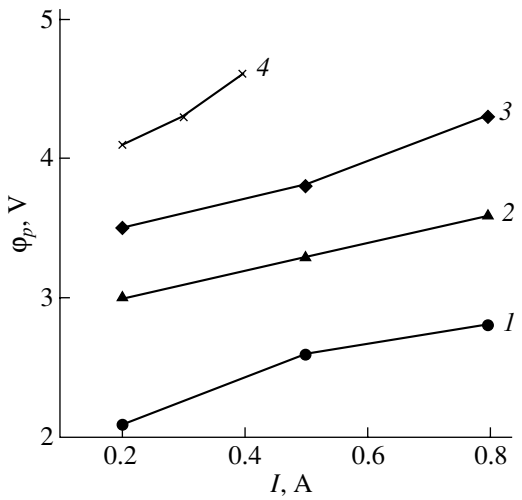


Fig. 4. Plasma potential φ_p as a function of the discharge current. The pressure in the vacuum chamber is (*1*) 8.0×10^{-2}, (*2*) 5.3×10^{-2}, (*3*) 2.7×10^{-2}, and (*4*) 0.8×10^{-2} Pa. The voltage between the electrodes of the second stage is 150 V.

tial increases by ~0.5 V. The error in measuring ϕ_p from the inflexion point of the emission probe characteristic is determined by the potential drop along the filament, the amplitude of the modulating voltage, and space-charge effects. To suppress these effects, we had to decrease both the ratio of the filament emission current to the electron saturation current from the plasma and the thickness of the heated filament [9]. Estimates showed that, in our experiments, the error in measuring ϕ_p was 1–2 V.

It follows from the results obtained that the ions are not confined electrostatically in the plasma of the second stage of the ion source; therefore, the ions produced in the discharge escape from the plasma to the screening electrode, the grid, and the anode. The ion current in the anode circuit is proportional to the loss area, which was 30–40 cm^2 under our experiment conditions. The area on the screening plate corresponding to the transverse dimensions of the nonmagnetized plasma region was nearly two times larger. The area of the anode hole in front of the grid was about 30 cm^2. Hence, the current of the ions arriving at the plate is, to a first approximation, about one-half of the total plasma ion current.

The model used in [12–14] is based on the assumption that the gas is ionized only by fast electrons. A fraction of these electrons arrive at the anode without energy loss in a time t_c and the other electrons relax over a time t_r, losing their energy in inelastic collisions with gas atoms (a part of this energy is spent on gas ionization). The characteristic ionization time is defined as

$$t_i = (n_0 \sigma_i v_e)^{-1}, \quad (1)$$

where n_0 is the neutral density, σ_i is the cross section for electron-impact ionization, and v_e is the velocity of fast electrons.

The relaxation time t_r is defined by a similar relationship in which the ionization cross section is replaced with the cross section σ_t for inelastic electron–neutral interactions. Using the above characteristic times, the ratio between the electron and ion currents can be represented in the form

$$I_e/I_i = t_i/t_c + t_i/t_r. \quad (2)$$

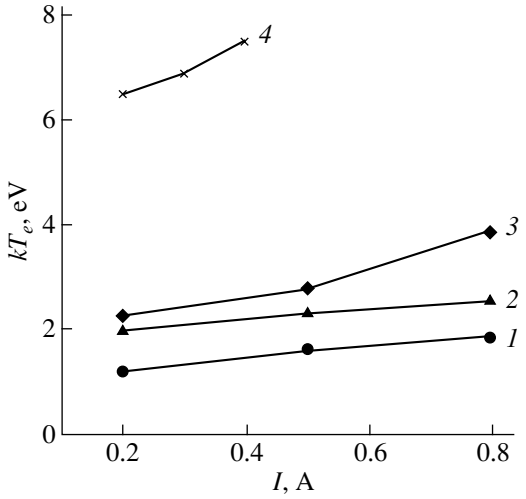


Fig. 5. Electron temperature kT_e as a function of the discharge current. The pressure in the vacuum chamber is (*1*) 8.0×10^{-2}, (*2*) 5.3×10^{-2}, (*3*) 2.7×10^{-2}, and (*4*) 0.8×10^{-2} Pa. The voltage between the electrodes of the second stage is 150 V.

It follows from this expression that the increase in the efficiency of ion extraction α with increasing pressure is caused by a decrease in the ratio t_i/t_c of the characteristic ionization time to the confinement time of fast electrons. According to [16], the ionization cross section σ_i and the cross section σ_t for the total inelastic loss in the 0.1- to 1-keV energy range depend on energy almost in the same manner and the ratio σ_i/σ_t usually lies in the range 0.2–0.5. Therefore, the second term in expression (2) only slightly influences the ratio I_e/I_i as the energy of fast electrons varies. In [14], it was shown that as the discharge voltage increases from 40 to 120 V (and, hence, the electron energy also increases), the effective loss area at the anode increases and the confinement time t_c of fast electrons decreases. A substantial increase in the efficiency of ion extraction observed in our experiments is explained by the fact that the product $\sigma_i v_e$ as a function of the electron energy increases rapidly to a value corresponding to the maximum of the cross section for electron-impact ionization. In the energy range under study, this cross section can be approximated by the formula [16]

$$\sigma_i = Z_W[10E_i(K - E_i)/[K(K + 8E_i)], \qquad (3)$$

where Z_W is the number of electrons in the outer shell, K is the electron kinetic energy, and E_i is the ionization energy.

In the electron energy range of 0.1–0.3 keV, the product $\sigma_i v_e$ varies only slightly; this should result in the saturation of the dependence of the efficiency of ion extraction on the electron energy.

To explain further increase in the efficiency of ion extraction, it should be taken into account that the model under consideration allows only for single ionizing collision events, whereas the increase in the energy of fast electrons enables them to make several such collisions. The maximum number of ionizing collisions can be estimated as $(eU/E_i)(\sigma_i v_e)/(\sigma_t v_e)$, where U is the voltage between the electrodes of the second stage (this voltage determines the energy of fast electrons). However, if the time during which electrons escape to the anode is too short for the total electron energy relaxation to occur, then the energy efficiency of the ion source, which is determined by the ratio of the ion-beam current to the power deposited in the discharge, decreases with increasing electron energy.

The effect of the diameter of the exit aperture of the hollow cathode on the efficiency of ion extraction (Fig. 2a) (which remains unexplained in the framework of the above simple model) can be explained by the nonuniformity of the spatial distribution of the neutrals and fast electrons in the second stage of the gas-discharge system. The pressure of the gas coming from the cathode cavity and expanding in the anode cavity is maximum near the grid. The directed flow of fast electrons enters this elevated-pressure region. As a result, the ionization frequency near the grid is higher than the average ionization frequency in the anode cavity. A decrease in the aperture diameter and the corresponding increase in the gas pressure and the density of the electron flow cause a more strong nonuniformity of the generated plasma. Since the ions produced near the grid can escape from the plasma to the grid or return along the magnetic filed lines to the adjacent part of the anode, a fraction of the ions arriving at the screening plate decreases with decreasing aperture diameter.

The results of measurements of the plasma potential with the help of an electrostatic spectrometer coincide qualitatively with the results of probe measurements. The position of the maximum of the energy spectrum of the ions extracted from the plasma corresponds to a positive (relative to the anode) plasma potential of a few electronvolts. The small full width at half-maximum (FWHM) of the longitudinal energy spectrum (6.5 eV) indicates that the plasma potential is constant in the ion generation region and that plasma instabilities are absent. As the pressure increases, the FWHM of the ion energy spectrum increases; most probably, this is a consequence of the charge exchange of ions in the screening-electrode sheath.

CONCLUSIONS

An increase in the multipole magnetic field in the anode stage of the discharge and the use of a multipole magnetic field in the hollow cathode of a glow discharge have made it possible to expand the operating pressure range of an ion source with a cold cathode to pressures as low as $\sim 5 \times 10^{-3}$ Pa.

The ratio of the current of the ions extracted from the plasma to the plasma-cathode current increases monotonically as the gas pressure and the voltage between the electrodes of the second stage of the discharge increase. The maximum ion current is close to the plasma-cathode current. As the pressure increases higher than $\sim 5 \times 10^{-2}$ Pa, the efficiency of ion extraction decreases.

The diameter of the cathode aperture has a substantial effect on the efficiency of ion extraction from the plasma; this is explained by a change in the spatial nonuniformity of the plasma produced.

The plasma potential in the anode stage of the discharge remains positive relative to the anode (1–5 V) as the current varies within the range 0.2–0.8 A, the gas pressure varies within the range 0.05–0.005 Pa, and the electron energy varies within the range 100–300 eV. The plasma electron temperature is 1–8 eV. The minimum FWHM of the energy spectrum of the ions extracted from the plasma at the minimum operating pressure is 6.5 eV.

REFERENCES

1. N. V. Gavrilov, G. A. Mesyats, G. V. Radkovskii, *et al.*, Surf. Coat. Technol. **96**, 81 (1997).

2. A. V. Vizir', E. M. Oks, P. M. Shchanin, and G. Yu. Yushkov, Zh. Tekh. Fiz. **67** (6), 28 (1997) [Tech. Phys. **42**, 611 (1997)].
3. N. V. Gavrilov and A. S. Kamenetskikh, Dokl. Akad. Nauk **394** (2), 1 (2004) [Dokl. Phys. **49**, 19 (2004)].
4. S. P. Nikitin, D. F. Chichigin, and P. V. Tretnikov, in *Proceedings of the 10th International Conference on Ion Sources, Dubna, 2003*, p. 215.
5. R. Limpaecher and K. R. MacKenzie, Rev. Sci. Instrum. **44**, 726 (1973).
6. N. V. Gavrilov, RF Patent No. 2134921 (1997).
7. K. N. Leung, N. Hershkowitz, and K. R. Mackenzie, Phys. Fluids **19**, 1045 (1976).
8. R. F. Kemp and J. M. Sellen, Rev. Sci. Instrum. **37**, 455 (1966).
9. J. R. Smith, N. Hershkowitz, and P. Coakley, Rev. Sci. Instrum. **50**, 210 (1979).
10. N. I. Ionov, Zh. Tekh. Fiz. **34**, 769 (1964) [Sov. Phys. Tech. Phys. **9**, 993 (1964)].
11. Yu. M. Kagan and V. I. Perel', Usp. Fiz. Nauk **81**, 409 (1963).
12. K. N. Leung, R. E. Kribel, A. P. Goede, *et al.*, Phys. Lett. A **66A**, 112 (1978).
13. A. J. T. Holmes, Rev. Sci. Instrum. **52**, 1814 (1981).
14. D. Cope and J. H. Keller, J. Appl. Phys. **56**, 96 (1984).
15. A. S. Metel', Zh. Tekh. Fiz. **54**, 241 (1984) [Sov. Phys. Tech. Phys. **29**, 141 (1984)].
16. V. E. Golant, A. P. Zhilinskiĭ, and I. E. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).

Translated by E. Satunina

Technical Physics, Vol. 49, No. 9, 2004, pp. 1208–1212. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 103–107.

ELECTRON AND ION BEAMS, ACCELERATORS

Kinetic Equation for a Relativistic Electron Beam Propagating along an External Magnetic Field in Dense and Rare Gas–Plasma Media

E. K. Kolesnikov and A. S. Manuĭlov

Smirnov Research Institute of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg, 198504 Russia

e-mail: Kolesnikov_evg@mail.ru

Received December 9, 2003

Abstract—A kinetic equation that describes the transverse dynamics of an axisymmetric paraxial relativistic electron beam propagating along an external magnetic field in a gas–plasma medium is derived with allowance for the influence of the self-consistent electromagnetic field on the beam, the effects related to the nonlaminar motion and rotation of the beam electrons at the exit from the injector, and the scattering and energy loss of the beam electrons in their collisions with the neutral particles of the background gas. © *2004 MAIK "Nauka/Interperiodica".*

1. FORMULATION OF THE KINETIC PROBLEM ABOUT THE EVOLUTION OF A RELATIVISTIC ELECTRON BEAM TRANSPORTED IN A SCATTERING GAS–PLASMA MEDIUM

New areas of application of relativistic electron beams (REBs) stimulate further investigations of the dynamics of their transport in gas–plasma media [1–24]. Since the transport of an REB in a gas–plasma medium is strongly nonequilibrium in character and is largely dominated by the collective electromagnetic field of the beam and plasma charges and currents, a natural methodological approach to constructing models of the transport of REBs in gas–plasma media is based on a mathematical apparatus involving the Vlasov–Boltzmann kinetic equations with self-consistent field and their consequences—the equations for the moments of the distribution function of the beam particles.

The kinetic equation describing the transverse dynamics of azimuthally symmetric paraxial REBs in dense gas–plasma media was first obtained by Lee [6].

The main goal of the present paper is to derive a kinetic equation that describes the transverse dynamics of paraxial REBs propagating along an external magnetic field in a dense gas–plasma medium and in a rare plasma in the ion focus regime.

In the kinetic description, the beam is characterized by the distribution function $\Phi_b(\mathbf{r}, \mathbf{p}, t)$, which has the meaning of the mathematical expectation of the number of relativistic electrons in a six-dimensional space whose six coordinates comprise the three spatial coordinates r and the three components of the relativistic momentum $\mathbf{p} = m_0\gamma\mathbf{v}$ (where m_0 is the rest mass of an electron, $\mathbf{v}$ is its velocity, and $\gamma = 1/(1 - v^2/c^2)^{1/2}$ is the Lorentz factor).

The evolution of the distribution function of the beam particles during the transport process is generally described by the Vlasov–Boltzmann kinetic equation [22–24]

$$\frac{\partial\Phi_b}{\partial t} + \mathbf{v}\cdot\nabla_\mathbf{r}\Phi_b + e\left[\mathbf{E} + \frac{1}{c}(\mathbf{v}\times\mathbf{B})\right]\nabla_\mathbf{p}\Phi_b = I_{sc}. \quad (1)$$

Here, $\mathbf{E}$ and $\mathbf{B}$ are the strength and induction of the electric and magnetic self-consistent fields, respectively, and the so-called collision integral I_{sc} on the right-hand side characterizes the effect of a change in the distribution function due to collisions of the beam particles with the neutral particles of the background gas.

Equation (1) should be solved together with Maxwell's equations for the self-consistent electromagnetic field:

$$\nabla\times\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}, \quad (2)$$

$$\nabla\times\mathbf{B} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}, \quad (3)$$

$$\nabla\cdot\mathbf{E} = 4\pi\rho, \quad (4)$$

$$\nabla\cdot\mathbf{B} = 0. \quad (5)$$

Here,

$$\rho = \rho_b + \rho_p, \quad \mathbf{J} = \mathbf{J}_b + \mathbf{J}_p, \quad (6)$$

1063-7842/04/4909-1208$26.00 © 2004 MAIK "Nauka/Interperiodica"

where the charge and current densities of the beam particles, ρ_b and $\mathbf{J}_b$, are determined by the corresponding moments of the distribution function Φ_b,

$$\rho_b = e\int \Phi_b(\mathbf{r}, \mathbf{p}, t)d\mathbf{p}, \quad (7)$$

$$\mathbf{J}_b = e\int \frac{\mathbf{p}}{\gamma m_0}\Phi_b(\mathbf{r}, \mathbf{p}, t)d\mathbf{p}; \quad (8)$$

and ρ_p and $\mathbf{J}_p$ are the densities of the charge and current that are induced in the plasma.

In what follows, we will restrict ourselves to studying some specific cases of transport of an REB in a gas–plasma medium, namely, those in which the charge and current densities of the plasma in the main part ("body") of the beam either can be treated as known functions of the coordinates or can be explicitly expressed in terms of the charge and current densities of the beam. Among such cases are the regimes of propagation of a charge- or a current-neutralized beam in a plasma in which the electron density n_Φ is much higher than the electron beam density n_b.

For a charge- and current-neutralized beam, we have

$$\rho_p = (1-\alpha_c)\rho_b, \quad \mathbf{J}_p = (1-\alpha_m)\mathbf{J}_b,$$

where α_c and α_m are the degrees of charge and current (magnetic) neutralization of the beam, respectively.

We will also consider the transport of a high-density pulsed REB in a rare plasma ($n_b \gg n_\Phi$) in the ion focus regime [9, 11–13]. The main feature of the ion focus regime is that the pressure of the background gas–plasma medium is low enough for the transverse component of the electric field excited by the front of the REB to expel the electrons produced in a preformed plasma channel from the beam region without producing additional ionization of the background plasma. This situation takes place under the condition [9]

$$\lambda_i \gg R_b,$$

where λ_i is the characteristic scale length on which the avalanche ionization develops and R_b is the characteristic beam radius.

In addition, the beam pulse duration τ_b should satisfy the condition $\tau_b \gg \tau_e$, where τ_e is the characteristic time during which the electric field of an REB expels the plasma electrons away from the beam region. Since the electrons of the background plasma are pushed away from the beam path by the beam electric field, the majority of the beam electrons propagate against the background of the plasma ions, which can be regarded as being immobile provided that the beam pulse duration satisfies the additional condition $\tau_b \ll \tau_i$ (where τ_i is the characteristic bounce period of the ions in the beam region) [9, 11].

In the above cases, Eqs. (1)–(5) with relationships (6)–(8) constitute a closed set of equations for the distribution function $\Phi_b(\mathbf{r}, \mathbf{p}, t)$ of the beam electrons and for the self-consistent electric field.

2. SEPARATE FORMULATION OF THE PROBLEMS OF THE LONGITUDINAL AND TRANSVERSE MOTIONS OF A BEAM PARTICLE IN THE PARAXIAL APPROXIMATION

The solution of the problem formulated in Section 1, i.e., the solution of Eqs. (1)–(5) supplemented with the corresponding initial and boundary conditions, is generally a very difficult task. In some particular cases, however, the solution can be substantially simplified by making additional assumptions about the properties of the beam. Thus, in the beam transport problems, the case of most practical interest is that of a so-called paraxial beam whose particles move at small angles relative to a certain axis (the z axis) and, accordingly, satisfy the condition

$$\frac{v_\perp}{v_z} \equiv \nu \ll 1, \quad (9)$$

where $v_\perp$ is the transverse (with respect to the z axis) velocity component of a beam particle.

Let us show that condition (9) greatly simplifies Eqs. (1)–(5). To do this, we consider the equation of motion of a single beam electron:

$$\frac{dp}{dt} = e\left[\mathbf{E} + \frac{1}{c}(\mathbf{v}\times\mathbf{B}^*)\right] + \mathbf{G}_{sc}. \quad (10)$$

Here, $\mathbf{B}^* = \mathbf{B} + \mathbf{B}_0$; $\mathbf{E}$ and $\mathbf{B}$ are the strength and induction of the electric and magnetic collective fields of the beam–plasma system, respectively; $\mathbf{B}_0 = B_0\mathbf{e}_z$ is the induction of the external magnetic field, which is assumed to be uniform and to be directed along the z axis; and $\mathbf{G}_{sc}$ is the fluctuating force due to collisions of the beam electron with the particles of the background medium.

We introduce the scalar and vector potentials of the electromagnetic field, φ and $\mathbf{A}^*$, which are expressed in terms of the vectors $\mathbf{E}$ and $\mathbf{B}^*$ via the well-known relationships

$$\mathbf{E} = -\nabla\varphi - \frac{1}{c}\frac{\partial\mathbf{A}^*}{\partial t}, \quad (11)$$

$$\mathbf{B}^* = \nabla\times\mathbf{A}^*, \quad (12)$$

where $\mathbf{A}^* = \mathbf{A}_0 + \mathbf{A}$, $\mathbf{A}_0$ is the vector potential of the external magnetic field, and φ and $\mathbf{A}$ are the potentials of the collective electromagnetic field.

Assuming that $B_0 = \text{const}$, we substitute expressions (11) and (12) into Maxwell's equations (2)–(5) and impose the Lorentz gauge condition ($\nabla\cdot\mathbf{A}^* + (1/c)\partial\varphi/\partial t) = 0$ to arrive at the following set of d'Alem-

bert's equations for the potentials φ and $\mathbf{A}$:

$$\Delta\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -4\pi\rho, \tag{13}$$

$$\Delta\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\frac{4\pi}{c}\mathbf{J}. \tag{14}$$

We treat the problem in a quasi-steady-state formulation under the assumption that the characteristic time scale on which the beam parameters vary, $\tau_b \approx f(\partial f/\partial t)^{-1}$, satisfies the condition

$$\tau_b > \frac{R_\perp}{c}\nu^{-\frac{1}{2}}, \tag{15}$$

where $R_\perp = f/|\nabla_\perp f|$ is the characteristic transverse scale of variations of the system parameters.

Taking into account the fact that, in the paraxial approximation, the longitudinal and transverse gradients of the quantities associated with the beam satisfy the relationship $|\partial f/\partial z/\nabla_\perp f| \approx \nu$, we see that, to first order in the parameter ν, Eqs. (13) and (14) can be written as

$$\Delta_\perp\varphi = -4\pi\rho + O(\nu), \tag{16}$$

$$\Delta_\perp\mathbf{A} = -\frac{4\pi}{c}\mathbf{J} + O(\nu), \tag{17}$$

where $\Delta_\perp$ is the Laplace operator with respect to the transverse coordinates.

We substitute expressions (11) and (12) into the equation of motion (10) and, in the resulting equation, ignore the motion in the z direction:

$$\frac{d}{dt}(m_0\gamma\mathbf{v}_\perp) = e\left[-\nabla_\perp\varphi - \frac{1}{c}\frac{\partial\mathbf{A}_\perp}{\partial t} + \frac{1}{c}(\mathbf{v}\times\nabla\times\mathbf{A}^*)_\perp\right] + (\mathbf{G}_{sc})_\perp. \tag{18}$$

An analysis shows that, to first order in the parameter ν, the parameters of the paraxial beam satisfy the relationship $1/c(\mathbf{v}\times\nabla_\perp\times\mathbf{A}^*) \approx v_z/c\nabla_\perp A_z + (\mathbf{v}_\perp/c) + \mathbf{i}_z B_0$ (where $\mathbf{i}_z$ is a unit vector directed along the z axis) and that, under quasi-steady state condition (15), the time-dependent term $1/c(\partial\mathbf{A}_\perp/\partial t)$ on the right-hand side of Eq. (18) can be ignored. Consequently, to within first-order terms, Eq. (18) can be reduced to the equation

$$\frac{d}{dt}m_0\gamma\mathbf{v}_\perp = e\left[-\nabla_\perp(\varphi - \beta A_z) + \frac{\mathbf{v}_\perp}{c} + \mathbf{i}_z B_0\right] + (\mathbf{G}_{sc})_\perp. \tag{19}$$

In the paraxial approximation, the law for the longitudinal motion of a beam particle can be determined from the energy equation

$$\frac{d\varepsilon}{dt} = \mathbf{v}\mathbf{F}, \tag{20}$$

which, in the case at hand, has the form

$$\frac{d\varepsilon}{dt} = e\mathbf{v}\mathbf{E} + \mathbf{v}\mathbf{G}_{sc} = ev_z E_z + e\mathbf{v}_\perp\mathbf{E}_\perp + \mathbf{v}\mathbf{G}_{sc}, \tag{21}$$

where $\varepsilon = m_0c^2\gamma$ is the relativistic energy of the particle.

Using relationship (11) to express E_z and $\mathbf{E}_\perp$ in terms of the potentials and taking into account conditions (9) and (15) (which, respectively, indicate that the beam is paraxial and is quasi-steady), we can show that the first term on the right-hand side of Eq. (21) is on the order of $\nu^{1/2}$, while the second term is on the order of ν. Therefore, to first order in the small parameter ν, the field terms on the right-hand side of Eq. (21) can be ignored. As a result, in the case under consideration, Eq. (21) may be written approximately as

$$\frac{d\varepsilon}{dt} = \mathbf{v}\mathbf{G}_{sc}. \tag{22}$$

Under conditions such that a beam particle loses an insignificant fraction of its energy in each collision with a background particle, the energy losses on time scales $\Delta t \gg \tau_{sc}$ (where τ_{sc} is the mean time between collisions) can be calculated by approximating the right-hand side of Eq. (22) by a continuous function $(-d\varepsilon/dt)_{sc}$, which is a known function of the energy of a beam particle for a given background medium. Hence, under the assumption that the energy losses are continuous, the energy equation can be written in the form

$$\frac{d\varepsilon}{dt} = \left(-\frac{d\varepsilon}{dt}\right)_{sc}. \tag{23}$$

Taking into account the fact that, in the paraxial approximation, the longitudinal component of the velocity of a beam particle has the form $v_z = v(1 + O(\nu^2))$ and that the velocity of the particle is related to its energy ε by the relationship $v = c(1 - m^2c^4/\varepsilon^2)^{1/2}$, we arrive at the equation

$$\frac{dz}{dt} = c(1 - m^2c^4/\varepsilon^2(t))^{1/2}. \tag{24}$$

We integrate Eq. (24) with the initial conditions $t = \tau$ and $\mathbf{z} = 0$ to obtain the following law for the longitudinal motion of a beam particle injected at the time τ:

$$z = c\int_\tau^t (1 - m^2c^4/\varepsilon^2(t'))^{\frac{1}{2}}dt', \tag{25}$$

where the dependence $\varepsilon(t)$ is determined from the solution to Eq. (23),

$$t - \tau = \int_{\varepsilon_0}^{\varepsilon}\frac{d\varepsilon'}{[-d\varepsilon(\varepsilon')/dt]_{sc}}, \tag{26}$$

with ε_0 being the initial particle energy at the exit from the injector.

It follows from Eqs. (25) and (26) that, in the approximation in question, the longitudinal motion of the beam particles is independent of their transverse motion and is deterministic in character.

Note that one of the assumptions made in deriving Eqs. (25) and (26), namely, the assumption that the energy losses are continuous, is generally correct only when the energy is lost predominantly in ionization and excitation of the atoms of the medium. For beam electrons of extremely high energies ($\varepsilon \geq \varepsilon_{cr}$, where ε_{cr} is the critical energy for a given background medium [23]), i.e., when the energy is lost mainly in the form of bremsstrahlung, the assumption that the energy losses are continuous may turn out to be incorrect because of strong statistical fluctuations in the energy the high-energy electrons lose by bremsstrahlung [23].

3. KINETIC EQUATION FOR THE DISTRIBUTION FUNCTION OF THE PARTICLES IN A SEGMENT OF A PARAXIAL BEAM PROPAGATING ALONG AN EXTERNAL MAGNETIC FIELD IN A SCATTERING GAS–PLASMA MEDIUM

The above analysis has shown that, in the paraxial approximation, the longitudinal motion of the beam particles is deterministic in character. In contrast to the longitudinal motion, the transverse dynamics of the beam particles is random, so that the beam state in the corresponding phase space can be properly described only statistically.

Let us partition the beam into thin transverse segments S^τ, each of which is injected at the time $t = \tau$ and contains a fixed number of particles.

We assume that the beam is monoenergetic at the exit from the injector and propagates through a uniform medium. Under the above assumptions, Eqs. (25) and (26) imply that all the particles in the segments S^τ move in the same manner in the z direction and, at any time, they have the same energy $\varepsilon(t)$ and the same relativistic mass $m^* = m\gamma = \varepsilon(t)/c^2$; moreover, during the beam propagation, the segments do not overlap.

For each of the segments S^τ, we introduce the function $f^\tau(\mathbf{r}_\perp, \mathbf{p}_\perp, t)$, which describes the distribution of the particles within the segment over the transverse coordinates $\mathbf{r}_\perp$ and transverse momenta $\mathbf{p}_\perp$. The evolution of the distribution function is described by the kinetic equation

$$\frac{\partial f^\tau}{\partial t} + \mathbf{v}_\perp \nabla_{\mathbf{r}_\perp} f^\tau + \mathbf{F}_\perp \nabla_{\mathbf{p}_\perp} f^\tau = I_{sc}, \qquad (27)$$

where $\mathbf{v}_\perp = d\mathbf{r}_\perp/dt = \mathbf{p}_\perp/m\gamma$, $\mathbf{F}_\perp$ is the transverse component of the force exerted by the self-consistent electromagnetic field on a beam particle, and I_{sc} is the collision integral.

Note that, in the case under analysis, Eq. (18) yields the following expression for the force $\mathbf{F}_\perp$:

$$\mathbf{F}_\perp = -e\nabla_\perp(\varphi - \beta A_z) + \Omega_b \mathbf{p}_\perp \times \mathbf{i}_z, \qquad (28)$$

where $\Omega_b = |e|B_0/(m\gamma c)$ is the gyrofrequency of the beam particles in the external magnetic field.

For conditions dominated by multiple small-angle elastic scatterings, the collision integral in Eq. (27) can be written in the form of the Fokker–Planck collision integral [22–24]:

$$I_{sc} = -\sum_\alpha \frac{\partial}{\partial p_{\perp\alpha}} A_\alpha f + \sum_{\alpha,\beta} \frac{\partial^2}{\partial p_{\perp\alpha}\partial p_{\perp\beta}} B_{\alpha\beta} f. \qquad (29)$$

Here, the Fokker–Planck coefficients are given by the expressions

$$A_\alpha = -\lim_{\Delta t \to 0} \frac{\langle \Delta p_{\perp\alpha} \rangle}{\Delta t}, \qquad (30)$$

$$B_{\alpha\beta} = \frac{1}{2}\lim_{\Delta t \to 0} \frac{\langle \Delta p_{\perp\alpha} \Delta p_{\perp\beta} \rangle}{\Delta t}, \qquad (31)$$

where

$$\langle \Delta p_{\perp\alpha} \rangle = \int w(\mathbf{p}_\perp, \Delta\mathbf{p}_\perp, \Delta t) \Delta p_{\perp\alpha} d\Delta\mathbf{p}, \qquad (32)$$

$$\langle \Delta p_{\perp\alpha} \Delta p_{\perp\beta} \rangle = \int w(\mathbf{p}_\perp, \Delta\mathbf{p}_\perp, \Delta t) \Delta p_{\perp\alpha} \Delta p_{\perp\beta} d\Delta\mathbf{p}, \qquad (33)$$

with $w(\mathbf{p}_\perp, \Delta\mathbf{p}_\perp, \Delta t)$ being the probability for the transverse momentum $\mathbf{p}_\perp$ to change (due to multiple scatterings) by the amount $\Delta\mathbf{p}_\perp$ during the time interval Δt.

The coefficients A_α form a vector describing the rate of transport of the transverse momentum $\mathbf{p}_\perp$. The coefficients $B_{\alpha\beta}$ are the elements of the generalized diffusion tensor. Lee [6] showed that, for an isotropic distribution, the coefficients A_α and the off-diagonal elements of the tensor $B_{\alpha\beta}$ are zero and the diagonal elements of the diffusion tensor have the form

$$B_{\alpha\alpha} = \frac{m\gamma}{2} S, \qquad (34)$$

where

$$S = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{\langle \Delta p_\perp^2 \rangle}{2m\gamma}.$$

In the case of multiple small-angle elastic scatterings, the function S is determined by the magnitude of the total momentum and is independent of the transverse momentum component $\mathbf{p}_\perp$ [25, 26]. Consequently, the assumption that the scattering process is isotropic and elastic allows us to substantially simplify collision integral (29):

$$I_{sc} = \frac{m\gamma S}{2} \Delta_{\mathbf{p}_\perp} f^\tau. \qquad (35)$$

We insert expression (28) for the force $\mathbf{F}_\perp$ and expression (35) for the collision integral into Eq. (27) and also take into account the fact that, in the transport regimes under consideration, the effective scalar potential of the transverse electric field satisfies the relation $\varphi - \beta A_z = \varphi_0 - \beta\mu A_z$, where φ_0 is the potential of the electric field of the neutralizing ion background and the constant μ is given by the formula

$$\mu = 1 - \frac{(1 - \alpha_c)}{\beta^2(1 - \alpha_m)} r \qquad (36)$$

As a result, we arrive at the following form of Eq. (27):

$$\frac{\partial f^\tau}{\partial t} + \frac{\mathbf{p}_\perp}{\gamma m}\nabla_{\mathbf{r}_\perp} f^\tau + [-e\nabla_\perp(\varphi_0 - \beta\mu A_z) + \Omega_b \mathbf{p}_\perp \times \mathbf{i}_z]\nabla_{\mathbf{p}_\perp} f^\tau = \frac{m\gamma S}{2}\Delta_{\mathbf{p}_\perp} f^\tau, \tag{37}$$

in which, in accordance with Eq. (17), the potential A_z satisfies the equation

$$\Delta_\perp A_z = -\frac{4\pi}{c}(1-\alpha_m)J_b, \tag{38}$$

and the potential φ_0 is to be regarded as the known solution to Poisson's equation

$$\Delta_\perp \varphi_0 = 4\pi e n_\Phi. \tag{39}$$

We introduce the radius R_c at which the self-consistent electromagnetic field of the background plasma is screened. In other words, we impose the boundary condition

$$\varphi|_{r \geq R_c} = A_z|_{r \geq R_c} \equiv 0. \tag{40}$$

The solution to Eq. (38) that satisfies boundary condition (40) has the form

$$A_z = -\frac{2}{c}J_b(1-\alpha_m)\int d\mathbf{r}'_\perp \ln\frac{|\mathbf{r}_\perp - \mathbf{r}'_\perp|}{R_c} \times \int d\mathbf{p}_\perp f^\tau(\mathbf{r}'_\perp, \mathbf{p}_\perp, t). \tag{41}$$

With allowance for relationship (41), Eq. (37) can be treated as an integrodifferential equation for the distribution function $f^\tau(\mathbf{r}_\perp, \mathbf{p}_\perp, t)$ of the particles in a segment. This equation should be supplemented with the initial condition

$$f^\tau(\mathbf{r}_\perp, \mathbf{p}_\perp, t)\big|_{t=\tau} = f_0(\mathbf{r}_\perp, \mathbf{p}_\perp, \tau), \tag{42}$$

where $f_0(\mathbf{r}_\perp, \mathbf{p}_\perp, t)$ is a given distribution function over transverse coordinates and transverse momenta that describes the beam particles at the exit from the injector.

The kinetic equation derived in the present paper can serve as a basis for numerical simulations of the transverse dynamics of paraxial REBs in gas–plasma media in the presence of an external longitudinal magnetic field. The consequences of this equation—the equations for the moments of the distribution function of the beam particles and for the phase-averaged quantities—can be used to construct simplified models of the transverse dynamics of an REB.

REFERENCES

1. A. N. Didenko, V. P. Grigor'ev, and Yu. P. Usov, *High-Power Electronic Beams and Their Applications* (Atomizdat, Moscow, 1977) [in Russian].
2. A. A. Rukhadze, L. S. Bogdankevich, S. E. Rosinskiĭ, *et al.*, *Physics of High-Current Relativistic Electron Beams* (Atomizdat, Moscow, 1980) [in Russian].
3. R. B. Miller, *Introduction to the Physics of Intense Charged Particle Beams* (Plenum, New York, 1982; Mir, Moscow, 1984).
4. J. D. Lawson, *The Physics of Charged-Particle Beams* (Clarendon Press, Oxford, 1977; Mir, Moscow, 1980).
5. R. C. Davidson, *Theory of Nonneutral Plasmas* (Benjamin, New York, 1974; Mir, Moscow, 1978).
6. E. P. Lee, Phys. Fluids **19**, 60 (1976).
7. E. P. Lee, Phys. Fluids **21**, 1327 (1978).
8. H. S. Uhm and M. Lampe, Phys. Fluids **23**, 1574 (1980).
9. H. L. Buchanan, Phys. Fluids **30**, 221 (1987).
10. E. K. Kolesnikov and A. D. Savkin, Pis'ma Zh. Tekh. Fiz. **20** (1), 54 (1994) [Tech. Phys. Lett. **20**, 26 (1994)].
11. E. K. Kolesnikov and A. S. Manuĭlov, Radiotekh. Élektron. (Moscow) **37**, 694 (1992).
12. E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **70** (5), 68 (2000) [Tech. Phys. **45**, 591 (2000)].
13. E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **70** (7), 127 (2000) [Tech. Phys. **45**, 939 (2000)].
14. R. F. Fernsler, S. P. Slinker, and R. F. Hubbard, Phys. Fluids B **3**, 2696 (1991).
15. E. R. Nadezhdin and G. A. Sorokin, Fiz. Plazmy **9**, 988 (1983) [Sov. J. Plasma Phys. **9**, 576 (1983)].
16. E. R. Nadezhdin and G. A. Sorokin, Fiz. Plazmy **14**, 619 (1988) [Sov. J. Plasma Phys. **14**, 365 (1988)].
17. E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **67** (7), 108 (1997) [Tech. Phys. **42**, 819 (1997)].
18. E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **67** (11), 62 (1997) [Tech. Phys. **42**, 1297 (1997)].
19. E. K. Kolesnikov and A. S. Manuĭlov, Radiotekh. Élektron. (Moscow) **44**, 1331 (1999).
20. A. S. Manuĭlov, Zh. Tekh. Fiz. **70** (1), 76 (2000) [Tech. Phys. **45**, 74 (2000)].
21. E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **67** (6), 69 (1997) [Tech. Phys. **42**, 648 (1997)].
22. R. Libov, *Introduction to the Theory of Kinetic Equations* (Wiley, New York, 1969; Mir, Moscow, 1974).
23. S. Chandrashekhar, Rev. Mod. Phys. **15** (1943); *Stochastic Problems in Physics and Astronomy* (GIIL, Moscow, 1947).
24. C. Cercignani, *Theory and Application of the Boltzmann Equation* (Am. Elsevier, New York, 1975; Mir, Moscow, 1978).
25. J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1975; Inostrannaya Literatura, Moscow, 1965).
26. B. Rossi and S. Olbert, *Introduction to the Physics of Space* (McGraw-Hill, New York, 1970; Atomizdat, Moscow, 1974).

Translated by O. Khadin

Technical Physics, Vol. 49, No. 9, 2004, pp. 1213–1220. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 108–115.

ELECTRON AND ION BEAMS, ACCELERATORS

Simulation of Methods for a Rapid Determination of the Energy Spectrum of Bremsstrahlung from Electron Accelerators

N. G. Mordasov, D. M. Ivashchenko, A. M. Chlenov, and A. A. Astakhov

Federal State Unitary Enterprise Research Institute for Instrument Making, Lytkarino, Moscow oblast, 140080 Russia

e-mail: ris@tsr.ru

Received January 8, 2004

Abstract—Semiempirical computational methods for determining the energy spectrum of bremsstrahlung from pulsed electron accelerators are proposed. Input information for employing these methods includes the energy spectrum of electrons in a pulse, the effective angle of their incidence on the target, and the parameters of the converter target. It is shown that the methods in question can be used in a rapid determination of bremsstrahlung spectra for an arbitrary angle of the escape of bremsstrahlung, including its escape in the forward direction. Among other things, this provides a solution to the problem of dosimetry in radiation tests of electron systems. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Semiempirical computational methods developed for determining angular, flux, and energy features of electrons in continuous cylindrical and tubular high-current beams [1, 2] are based on measuring the properties of bremsstrahlung downstream of the converter target. These methods created necessary preconditions for employing numerical methods for determining the energy spectra of bremsstrahlung in a radiation pulse for known parameters of the converter target. In the context of the problem that is addressed here and which is associated with exploring and identifying the bremsstrahlung fields in simulating facilities for large-scale radiation tests of electron systems by monitoring the features of the field during radiation tests, the energy spectrum of bremsstrahlung within any interval of radiation-pulse duration is one of the most important properties of the bremsstrahlung field and is simultaneously an operative factor. The intensity of bremsstrahlung is also an important parameter of the operative factor. In calculating the spectral distributions of bremsstrahlung in a radiation pulse in the sample-irradiation plane, the form of the description of the cross section for bremsstrahlung production by electrons plays an important role. It is hardly possible to perform an exact theoretical analysis of the formation of a bremsstrahlung photon in the screened field of a nucleus of converter-target matter. The problem is that the respective Dirac wave equation cannot be solved in a finite form, since the wave functions involved are represented in the form of an infinite series [3]. At the present time, there is a large number of studies aimed at deriving analytic expressions for the bremsstrahlung-production cross section on the basis of various approximations [4, 5]. The best of the currently available versions of the analytic representation for the bremsstrahlung-production cross section [5] is characterized by an estimated error of not more than 5%, which is thought to be quite sufficient for solving a wide range of applied problems. The spectral distribution of bremsstrahlung photons originating from thick targets (whose thickness is larger than the total range of primary electrons) is calculated with allowance for numerous processes of electron interaction with matter. For this reason, the volume of the calculations that would ensure an acceptable accuracy of the results (not poorer than 15%) appears to be quite large [6]. Such calculations, which are based on the Monte Carlo method, consume much machine time even if one employs a powerful PC. This generates inconveniences and rules out the possibility of an on-line derivation of information about the energy spectrum of bremsstrahlung between the pulses of radiation from simulating facilities in performing standard radiation tests of electron equipment. Computational procedures known as semiempirical computational methods are developed for such problems. Such procedures are based on breaking down the converter target into "thin" layers and on employing the mean features of an electron beam in each layer [7]. The measurements performed within the computational procedure from [8] take into account the ring shape of the beam cross section, the angular distribution of electrons entering the converter target, specific physical parameters of "elementary" target layers, and so on, making it possible to develop an express method for determining the energy spectrum of bremsstrahlung from accelerators producing electron beams of various shapes [9].

1063-7842/04/4909-1213$26.00

METHODOLOGICAL COMPUTATIONAL PART

Let us consider the most general scheme of our semiempirical computational method. A beam formed by monoenergetic electrons of kinetic energy E is normally incident on a target manufactured from a material characterized by a mass number A, a charge number Z, and a density ρ. We break down the target, whose thickness is denoted by D, into thin layers (see Fig. 1). We will dwell at greater length on the thickness of an "elementary" layer somewhat later. For the time being, we only require that its thickness be such that the electron energy loss in a layer is much less than the primary kinetic energy E.

We denote by Δd the elementary-layer thickness and by M the number of elementary layers in the target and analyze bremsstrahlung-photon production in an arbitrary ith layer of the target. The depth at which the median plane of this layer lays in the target is

$$d_i = \Delta d(i - 0.5). \tag{1}$$

Further, we assume that all electrons that have reached the ith layer have an identical energy E_i; however, the number of these electrons have decreased. The fraction of electrons that have reached the ith layer can be determined by using empirical equations that relate the number of electrons to the depth of their penetration into the target (electron transmission). Because of scattering processes, the velocities of the electrons in the ith layer have different directions; we denote by $U(\Theta, E_i)$ the respective angular distribution. By using the above notation, we can specify the probability for an electron to occur in the ith layer, to undergo scattering in it into the solid angle $\sin\Theta d\Theta d\varphi$, and to produce a photon whose energy lies in the interval between E_γ and $E_\gamma + dE_\gamma$ and whose emission angle with respect to the electron-momentum direction is ω. We have

$$\left(\frac{d^2\gamma}{dE_\gamma d\Omega^1}\right)_i = (N_{\text{ef}})_i \tau_i U_i(\Theta, E_i) \times \frac{d^2\sigma}{dE_\gamma d\Omega}(E_i, E_\gamma, \omega)\sin\Theta d\Theta d\varphi, \tag{2}$$

where $(N_{\text{ef}})_i$ is the effective number of target atoms in the layer (it depends on the angular distribution of electrons),

$$\frac{d^2\sigma}{dE_\gamma d\Omega}(E_i, E_\gamma, \omega)$$

is the differential cross section for bremsstrahlung production with respect to the emission angle and photon energy, and τ_i is the probability for an electron to reach the ith layer.

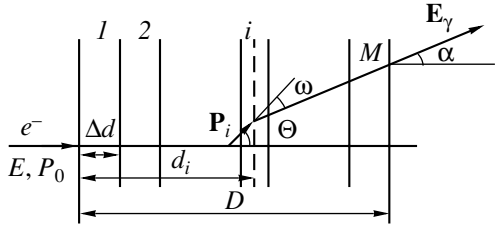


Fig. 1. Scheme of photon production in the target.

The angle α at which a photon escapes from the target is related to the azimuthal and polar electron-scattering angles Θ and φ and the angle ω between the electron and bremsstrahlung-photon momenta by the well-known equation

$$\cos\omega = \cos\alpha\cos\Theta + \sin\alpha\sin\Theta\cos\varphi. \tag{3}$$

Photons produced in the interior of the target are absorbed in its subsequent layers. This process is taken into account by introducing the coefficient η_i of bremsstrahlung self-absorption in the form

$$\eta_i(E_\gamma, \alpha) = \exp\left[-\mu(E_\gamma)\frac{D - d_i}{\cos\alpha}\right], \tag{4}$$

where $\mu(E_\gamma)$ is the mass coefficient of photon absorption in a target material.

In order to obtain the total number of bremsstrahlung photons of energy between E_γ and $E_\gamma + dE_\gamma$ that were emitted at an angle α from the target into the solid angle $d\Omega^1$, it is therefore necessary to integrate expression (2) over the entire region of the angles Θ and φ, to take into account bremsstrahlung self-absorption, and to perform summation over all target layers. Eventually, we have

$$\frac{d^2Y}{dE_\gamma d\Omega^1} = \sum_{i=1}^{M} \tau_i (N_{\text{ef}})_i \eta_i(E_\gamma, \alpha) \times \int_0^{2\pi}\int_0^{\pi} U(\Theta, E_i)\frac{d^2\sigma}{dE_\gamma d\Omega}(E_i, E_\gamma, \omega)\sin\Theta d\Theta d\varphi. \tag{5}$$

An implementation of the semiempirical method employing expression (5) requires refining the method for calculating all of the quantities involved in this expression, but it should be emphasized that the form of the angular distribution of electrons and the representation of the cross section for bremsstrahlung production are of prime importance here.

Among the known representations of the above double-differential cross section for bremsstrahlung pro-

duction, the Schiff formula [3]

$$\frac{d^2\sigma}{dE_\gamma dy} = \frac{4Z^2r_0^2y}{137E_\gamma}\left\{\frac{16y^2E}{(y^2+1)^4E_0} - \frac{(E_0+E)^2}{(y^2+1)^4E_0^2}\right.$$
$$\left.\times\left[\frac{E_0^2+E^2}{(y^2+1)^2E_0^2} - \frac{4y^2E}{(y^2+1)^4E_0}\right]\ln M(y)\right\}, \quad (6)$$

where $y = E_0\Theta_0$ is the reduced photon emission angle,

$$\frac{1}{M(y)} = \left(\frac{E_\gamma}{2EE_0}\right)^2 + \left[\frac{Z^{1/3}}{111(y^2+1)}\right]^2,$$

E_0 is the total energy, E is the scattered-electron energy, E_γ is the photon energy, and r_0 is the classical electron radius, is characterized by the maximum degree of universality and is relatively simple over the electron energy range 1–15 MeV.

In order to describe the angular distribution of electrons in the region of small angles, use is made of the first two terms of the Molière distribution [10],

$$U_i(\Theta)\Theta d\Theta = \frac{1}{2\pi}\left[2\exp(-\vartheta^2) + \frac{F_1(\vartheta)}{B}\right]d\vartheta; \quad (7)$$

where $F_1(\vartheta)$ is the tabulated Molière function, B is a parameter that increases slowly with increasing number of electron collisions, and

$$\vartheta = \frac{\Theta(E_0+mc^2)}{\sqrt{d_iB}}$$

is the reduced angle.

The first two terms of the distribution are quite sufficient here since the inclusion of the third term changes the total value by not more than 10% for the least favorable cases. For large electron-scattering angles ($\vartheta > 2.8$), we employed the asymptotic formula

$$U_i(\Theta)\Theta d\Theta = \frac{2d\vartheta}{B\vartheta^3\left\{1 - 4\vartheta^{-2}\left[1 + 2B^{-1}\ln\left(\frac{2\vartheta}{5}\right)\right]\right\}}, \quad (8)$$

which was proposed by Bethe [11].

With increasing depth of electron penetration into the target, the angular distribution of electrons broadens steadily, approaching an isotropic distribution at large depths [8].

Let us examine the problem of determining the effective number of target atoms in an elementary layer. As a rule, the effective number of atoms in a layer was determined according to [7, 8]—that is, with the aid of the relation

$$(N_{ef})_i = \frac{\Delta dN_A}{A\cos[(\langle\Theta_e^2\rangle)_i^{1/2}]}, \quad (9)$$

where Δd is the layer thickness in g cm^{-2} units,

$$(\langle\Theta_e^2\rangle)_i = \frac{d_iB}{(E_0+mc^2)^2}$$

is the mean square electron-scattering angle in a layer, and N_A is Avogadro's number.

But in fact, the path traveled by an electron scattered at an angle Θ in a layer is $\Delta d/\cos\Theta$, whence it follows that, in order to deduce the effective number of atoms, it is necessary to determine the quantity $\langle(\cos\Theta)^{-1}\rangle$ for the corresponding effective angular distribution of electrons. Since the code for respective calculations would become very cumbersome upon supplementing the computational procedure with the calculation of the effective number of atoms in a layer by evaluating the mean electron path for the rather complicated Molière distributions, a correction R_s was introduced in relation (9). This correction must take into account the distinction between exact relations and formula (9). The effective number of atoms was represented in the form

$$(N_{ef})_i = \frac{\Delta dN_A}{A\cos[(R_s\langle\Theta_e^2\rangle)_i^{1/2}]}. \quad (10)$$

The correction R_s was preliminarily estimated for Gaussian distributions, but it was ultimately determined only upon a thorough comparison with data from the literature on the spectra and angular distributions of bremsstrahlung. For the electron-energy range being considered, the best results were obtained at $R_s = 1.5$. Table 1, in which the mean paths for various values of the mean-square scattering angle according to calculations based on relations (9) and (10) are given along with $\langle(\cos\Theta)^{-1}\rangle$ values for the Gaussian representation of the angular distribution of electrons (first term of the Molière distribution), is presented here to illustrate the effect of the chosen correction on the mean path of an electron in a layer.

From an analysis of the data in Table 1, it follows that the mean values of the electron path in a layer without the above correction are smaller than the corresponding values determined for the first term of the Molière distribution and that the inclusion of one more term in the Molière distribution and the use of the Bethe asymptotic expression must lead to an increase in the mean path. In this respect, the mean-path values

Table 1. Mean path of an electron in an "elementary" layer

$\langle\Theta_e^2\rangle$	Formula (9)	Formula (10) at $R_s = 1.5$	$\langle(\cos\Theta)^{-1}\rangle$ for a Gaussian distribution
0.1	1.052	1.079	1.054
0.3	1.171	1.277	1.205
0.5	1.315	1.544	1.474
1	1.85	2.948	2.269

obtained with allowance for the correction R_s provide quite a good approximation to the most probable values of the mean path traveled by an electron in an elementary layer. From the table, it also follows that the mean path of electrons in an elementary layer increases quickly with increasing depth of electron penetration into the target ($\langle\Theta_e^2\rangle_i \sim d_i$), and this may lead to physically incorrect results. It is rather difficult to take exactly into account the effect of scattering processes on the mean path of an electron in deep layers, and this is beyond the scope of the developed procedure. In order to avoid physically meaningless results, the mean path traveled by an electron in an elementary layer was required to satisfy the condition $\cos[(\langle\Theta_e^2\rangle)_i R_s]^{1/2} \geq 0.5$; if $\cos[(\langle\Theta_e^2\rangle)_i R_s]^{1/2}$ fell below 0.5, the mean path of an electron in an elementary layer was taken to be $2\Delta d$. This limit was found empirically.

In our procedure, it is of importance to correctly take into account the electron-energy loss in a target material. In order to avoid encumbering the computational code with tables of the electron-energy loss, we made use of simple relations for mean electron-energy losses by ionization and radiation [12] with allowance for empirical corrections that make it possible to arrive at results similar to the computational data obtained in [13]. The resulting relation has the form

$$\frac{dE}{dx} = \left(\frac{dE}{dx}\right)_0 \frac{1}{1 + 0.02(E_0 - 10)}, \qquad (11)$$

where $(dE/dx)_0$ is the energy loss given by the formulas from [12].

The electron-transmission coefficient in terms of the number of particles is determined by the formulas presented by Tabato and Ito [14]; that is,

$$\tau_i = [1 + \exp(-S_0)]/\{1 \exp[(S_0 + 2)(d_i/R_{ex}) - S_0]\},$$

$$S_0 = a_1 \exp[a_2/(1 + a_3 E_{0m}^{a_4})], \qquad (12)$$

$$a_1 = b_1/Z^{b_2}, \quad a_2 = b_3 Z^{b_4}, \quad a_4 = b_5, \quad a_4 = b_6,$$

where E_{0m} is the kinetic energy of electrons in mc^2 units.

For the constants b_i (i = 1, 2, …, 6), those authors give the values of $b_1 = 10.63$, $b_2 = 0.232$, $b_3 = 0.22$, $b_4 = 0.462$, $b_5 = 0.042$, and $b_6 = 1.86$. For the extrapolated electron range R_{ex}, we employed the relation [14]

$$R_{ex} = c_1[(1/c_2)\ln(1 + c_2 E_{0m}) - c_3 E_{0m}/(1 + c_4 E_{0m}^{c_5})], \qquad (13)$$

where $c_1 = 0.2325A/Z^{1.209}$ g cm^{-2}, $c_2 = 0.000178Z$, $c_3 = 0.9891–0.000301Z$, $c_4 = 1.468–0.0118Z$, and $c_5 = 1.232/Z^{0.109}$.

In [7], the thickness of an elementary target layer was set to $0.001X_0$ (where X_0 is the radiation-length unit), irrespective of the electron energy, although the upper limit of the elementary-layer thickness is determined from the condition $\Delta d[dE/dx(E_0)] \ll E_0$, where $dE/dx(E_0)$ is the total electron-energy loss; therefore, the elementary-layer thickness can be increased for higher electron energies. With the aim of reducing the machine time consumed by the code, we tried to find greater values of the thickness and, on the basis of experimental results, chose the effective-thickness representation of the form

$$\Delta d = (0.005X_0 + E_0 0.0005X_0)/\cos Q, \qquad (14)$$

where Q is the effective angle of electron incidence on the target.

Specifically, we employed the thickness $0.1\Delta d$ for the first ten layers of the target and the thickness $0.2\Delta d$ for the second ten layers. This way of specifying the elementary-layer thickness is weakly sensitive to a further partition of layers and, at the same time, ensures a significant economy of the machine time in calculating the energy spectrum of bremsstrahlung. Since the stage of calculating the integral of the convolution of the double-differential cross section for bremsstrahlung production with the angular distribution of electrons determines the general rate of the program performance, we used the change of variables

$$U = \exp(-\Theta^2/\langle\Theta_e^2\rangle_i). \qquad (15)$$

In performing radiation-resistance tests of electron-equipment units, it is of particular interest to determine the energy spectrum of bremsstrahlung in the tested-sample plane, which is usually normal to the axis of electron-beam transportation; that is, it is necessary in this case to find the bremsstrahlung spectrum in a narrow angular range around the forward direction. In calculating the bremsstrahlung spectrum by the empirical computational method developed in [9], one can then introduce a number of additional simplifications that reduce the volume of calculations substantially. For the forward direction, Eq. (3), which relates the photon emission angle to the electron scattering angle, is significantly simplified. Substituting $\alpha = 0$ into it, we obtain $\cos\Theta = \cos\omega$; that is, $\omega = \Theta$. The bulk of the spectrum of bremsstrahlung in the forward direction is formed in the first layers of the target. The contribution of photons to the total bremsstrahlung intensity as a function of the depth of their production in the target is shown in Fig. 2 according to the calculation based on the results presented in [3]. At small depths in the target, the angular distribution of electrons has a pencil-like character, so that the bremsstrahlung spectrum is formed under the condition that the angle between the photon and electron momenta is small ($E \geq 2$ MeV); therefore, one can disregard the dependence of the shape of the energy distribution of bremsstrahlung on the bremsstrahlung-photon emission angle. Thus, it is

legitimate to factorize the double-differential cross section for bremsstrahlung production in the form

$$\frac{d^2\sigma}{dE_\gamma d\Omega} = \frac{d\sigma}{dE_\gamma}\frac{d\sigma}{d\Omega}, \tag{16}$$

where the factor $d\sigma/dE_\gamma$ specifies the dependence of the cross section on the bremsstrahlung energy exclusively and the factor $d\sigma/d\Omega$ determines the angular dependence of the cross section.

The general expression for the cross section can be represented in the form

$$\left(\frac{dY}{dE_\gamma}\right)_{\alpha=0} = 2\pi\sum_{i=1}^{M}\tau_i(N_{\text{ef}})_i\eta_i(E_\gamma)\frac{d\sigma}{dE_\gamma}(E_i, E_\gamma) \times\int_0^\pi U_i(\Theta)\frac{d\sigma}{d\Theta}(E_i, \Theta)\sin\Theta d\Theta. \tag{17}$$

For the angular distribution of bremsstrahlung, we selected the approximate expression

$$\frac{d\sigma}{d\Theta}(E_i, \Theta) = \frac{1}{2\pi\langle\Theta_b^2\rangle_i}e^{-\frac{\Theta^2}{\langle\Theta_b^2\rangle_i}} + \frac{1}{2\pi(10\langle\Theta_b^2\rangle_i)}e^{-\frac{\Theta^2}{10\langle\Theta_b^2\rangle_i}}, \tag{18}$$

where $(\langle\Theta_e^2\rangle_i)^{1/2} = mc^2/(E_i + mc^2)$.

Expression (18) describes well the true angular distribution of bremsstrahlung photons for the region where the spectrum of bremsstrahlung in the forward direction is formed. Moreover, the use of the angular distribution of bremsstrahlung photons in this form makes it possible to integrate expression (18) readily in the approximation of small angles—that is, in the approximation where the first term of the Molière distribution is used for the angular distribution of electrons and where $\sin\Theta$ is replaced by Θ. The result is

$$\left(\frac{dY}{dE_\gamma}\right)_{\alpha=0} = \sum_{i=1}^{M}\tau_i(N_{\text{ef}})_i\eta_i(E_\gamma)\frac{d\sigma}{dE_\gamma}(E_\gamma) \times\frac{11(\langle\Theta_b^2\rangle_i) + 2(\langle\Theta_e^2\rangle_i)}{10(\langle\Theta_b^2\rangle)_i^2 + 11(\langle\Theta_b^2\rangle)_i(\langle\Theta_e^2\rangle)_i + (\langle\Theta_e^2\rangle)_i^2}, \tag{19}$$

where

$$(\langle\Theta_e^2\rangle)_i = \frac{BR_T\sqrt{d}}{(E_0 + mc^2)}, \quad R_T = 1.02.$$

The differential cross section for bremsstrahlung production was taken from [3] for various values of the screening parameter γ. Specifically, we considered the

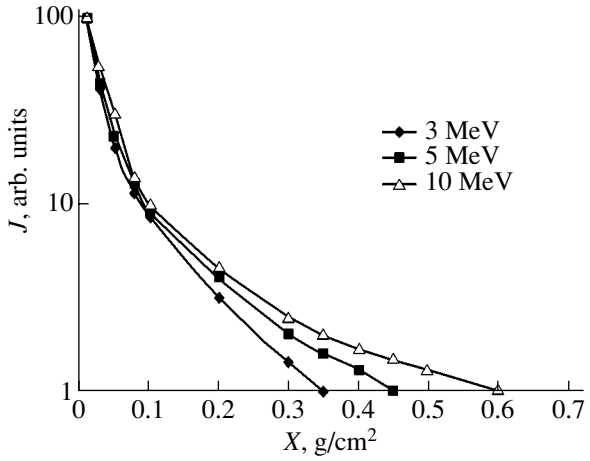


Fig. 2. Contribution of the tantalum-target thickness to the bremsstrahlung spectrum (J is the bremsstrahlung intensity).

case of $\gamma = 0$ (full screening),

$$d\sigma(E_\gamma) = \frac{4Z^2r_0^2dE_\gamma}{137E_\gamma}\left\{\left[1 + \left(\frac{E}{E_0}\right)^2 - \frac{2E}{3E_0}\right] \times\ln(183Z^{-1/3}) + \frac{E}{9E_0}\right\}; \tag{20}$$

the case of $\gamma < 2$ (intermediate screening),

$$d\sigma(E_\gamma) = \frac{4Z^2r_0^2dE_\gamma}{137E_\gamma}\left\{\left[1 + \frac{E^2}{E_0^2}\right]\left[\frac{\varphi_1(\gamma)}{4} - \frac{\ln Z}{3}\right] - \frac{2E}{3E_0}\left[\frac{\varphi_2(\gamma)}{4} - \frac{1}{3}\ln Z\right]\right\}; \tag{21}$$

and the case of $2 \le \gamma \le 15$ (intermediate screening),

$$d\sigma(E_\gamma) = \frac{4Z^2r_0^2dE_\gamma}{137E_\gamma}\left[1 + \left(\frac{E}{E_0}\right)^2 - \frac{2E}{3E_0}\right] \times\left[\ln\frac{2EE_0}{E_\gamma} - \frac{1}{2} - c(\gamma)\right]. \tag{22}$$

$$\varphi_1(\gamma) = 19.24 - 4\ln\left(\gamma + \frac{2}{\gamma+3}\right) - 0.12\gamma\exp\left(-\frac{\gamma}{3}\right). \tag{23}$$

The screening functions were taken in an analytic (rather in a graphical) form with allowance for a Coulomb correction according to the results presented in [15] and were determined with an error not exceeding 0.5%:

$$\varphi_2(\gamma) = \varphi_1(\gamma) - 0.027 - (0.8 - \gamma)^2, \quad \gamma \le 0.8, \qquad \varphi_2(\gamma) = \varphi_1(\gamma), \quad \gamma \ge 0.8. \tag{24}$$

The dependence of the function c on the screening parameter γ was taken from [3]. For the remaining factors in expression (19) $\{(N_{ef})_i, \tau_i, \eta_i, dE/dx\}$, we employed the same formulas as in calculating bremsstrahlung spectra for an arbitrary photon emission angle. Since the volume of the calculations decreased considerably under the forward-direction conditions, the thickness Δd of elementary target layers was chosen to be

$$\Delta d = 0.0001 X_0 (E_0 + 10), \tag{25}$$

where X_0 is the radiation-length unit, in order to simplify the computational code still further.

The above methodological framework is optimal in calculating the fields of bremsstrahlung from targets characterized by a high charge number ($70 \le Z \le 85$). The use of total-absorption targets from such materials (tantalum, tungsten, etc.) is not always economically justified. By way of example, we indicate that, in a prompt-action accelerator of the UIN-10 type, a target becomes disabled after one to three pulses. Therefore, it is often economically wiser to use targets from cheaper and lighter materials. The results of the experiments reported in [2] revealed that, if use is made of targets from iron ($Z = 26$), the intensity of bremsstrahlung decreases by not more than 30% over the entire energy range in relation to the corresponding intensity for a target from tantalum.

In going over to light-atom targets, the determination of the thickness of elementary layers into which the converter target is broken down must be refined. The range of applicability of multiple-scattering theory determines the maximum value of the elementary-layer thickness approximately. As was shown in [6], the number n of collisions in a material characterized by a charge number Z and a mass number A is given by

$$n = 10^{8.215} \frac{\Delta d \varphi^2}{A Z^{2/3} (1.13 + 3.77\phi^2)}, \tag{26}$$
$$\phi = Z/(137\beta),$$

where Δd is the depth of electron penetration into the target material and β is the electron speed in units of the speed of light.

Table 2. Layer thickness for various metals

Z	Δd, g cm^{-2}	$\Delta d \times 10^3$
4	0.00368	0.0538
13	0.00236	0.0988
29	0.00212	0.1656
42	0.00222	0.23
74	0.00298	0.472
82	0.0033	0.569

For various materials, the values calculated by formula (26) for the thickness of a layer where a relativistic electron ($\beta \approx 1$) undergoes 20 collisions are given in Table 2.

As can be seen from Table 2, the layer thickness does not change strongly with the charge number of the material. In considering bremsstrahlung-production processes, however, it is common practice to make use of units of the radiation length X_0, but, in these units, the layer thickness is a monotonically increasing function of the target charge number, this function changing approximately by an order of magnitude upon going over from beryllium to lead. In order to calculate the elementary-layer thickness in various materials, it is therefore proposed to use the relation

$$\Delta d(Z) = \frac{(Z + 10)(E_0 + 10)}{84 \cos Q} 0.0001 X_0, \tag{27}$$

where Q is the effective angle of electron incidence on the target and E_0 is the total electron energy.

The scheme according to which the thickness of an elementary layer changes versus its ordinal number is identical to that in calculating the spectra of bremsstrahlung in an arbitrary direction. Formula (25) leads to a charge-number dependence that is weaker than that in Table 2, but this is quite admissible if one considers that, in each collision, the mean electron scattering angle is smaller in light- than in heavy-atom materials. The resulting angular distributions and yields of bremsstrahlung photons are in good agreement with experimental data from [3]. On the basis of the above approaches, we developed a code for computing the energy spectrum of bremsstrahlung in the forward direction (LUE). The code makes it possible to compute the energy spectra of bremsstrahlung photons in the forward direction that are generated by electrons whose energy spectrum has the form of a set of individual lines. The thickness and the charge number of a single-layer target can have arbitrary values.

EXPERIMENTAL PART

The code for computing the bremsstrahlung spectrum for an arbitrary photon emission angle (ANGLE1) is written in the FORTRAN language. The input data for the code include target parameters, the angle at which electrons enter the target, the angle at which photons escape from it, and the electron energy. As the output information, the code provides the bremsstrahlung spectrum in photon (MeV sr electron)$^{-1}$ units and the total intensity. Calculations at individual stages (that is, calculations of the extrapolated electron range, of the electron-transmission coefficient, of the electron-energy loss, of photon self-absorption, and of the integral of a spatial convolution of the double-differential cross section for bremsstrahlung production with the angular distribution of electrons) are arranged as subroutines. For the operation of the code for integration to

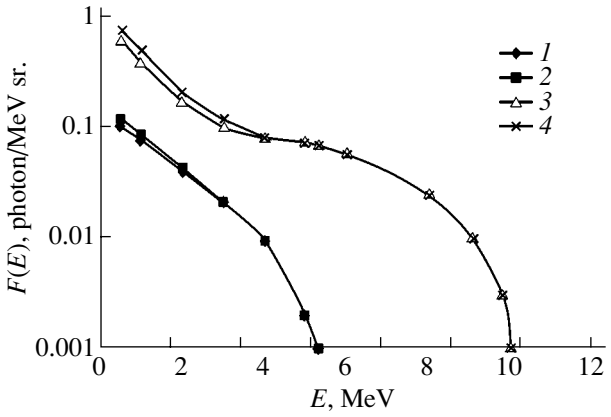


Fig. 3. Energy spectra of bremsstrahlung in the forward direction.

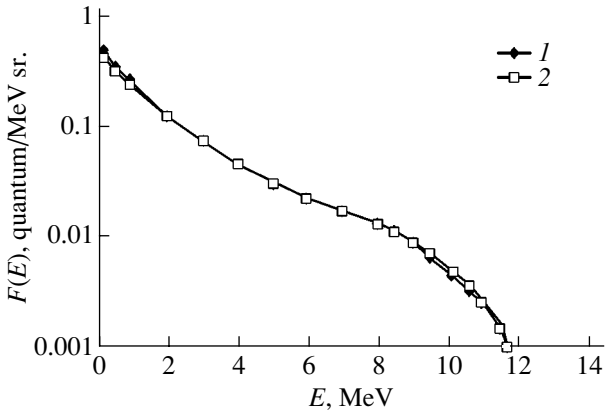


Fig. 4. Energy spectrum of bremsstrahlung from a LUI-10 accelerator: (*1*) Monte Carlo results and (*2*) ANGLE11 results.

be stable, the preset accuracy must not be higher than 0.05. Integration is performed by the Gauss method, and the accuracy in calculating the integral in question is an input parameter of the code. On average, it takes 3 s to compute one bremsstrahlung spectrum at an electron energy of 10 MeV.

Thus, the time it takes for a fourth-generation PC to obtain full information about the bremsstrahlung field by calculating, for example, the spectra for six values of the photon emission angle is 20 s. In calculations by the Monte Carlo method, the time required for accumulating 10^5 histories of electrons under the same conditions (these statistics are not very good for large photon emission angles) is an order of magnitude longer.

The results obtained by calculating the spectra of bremsstrahlung in the forward direction on the basis of simplifying the semiempirical procedure for an arbitrary photon emission angle were compared with experimental data from [17] for the electron energies of 5 and 10 MeV. These results are given in Fig. 3 for a tungsten target of thickness $D = 0.735$ g cm^{-2}. From this figure, it can be seen that, for the electron energies being considered, our results are in good agreement with data from the literature within the experimental errors. Thus, this comparison with the data from the literature gives sufficient grounds to believe that the results obtained on the basis of the refined semiempirical procedure employing the LUE code are quite reliable. In the calculations based on the LUE code, use is made of some subroutines also entering into the composition of the ANGLE1 code. The LUE code is written in the FORTRAN language, and the results of the calculations are represented in the form of a table that contains bremsstrahlung spectra in photon (MeV sr electron)$^{-1}$ units.

The energy spectrum of bremsstrahlung from a linear electron accelerator of the LUI-10 type (tubular electron beam) in the forward direction is displayed in Fig. 4 according to calculations by the Monte Carlo method [6] and by the above semiempirical method (ANGLE1 code) for a tantalum converter target that were based on the results obtained by determining the energy spectrum of electrons in a pulse with the aid of the procedure proposed in [1].

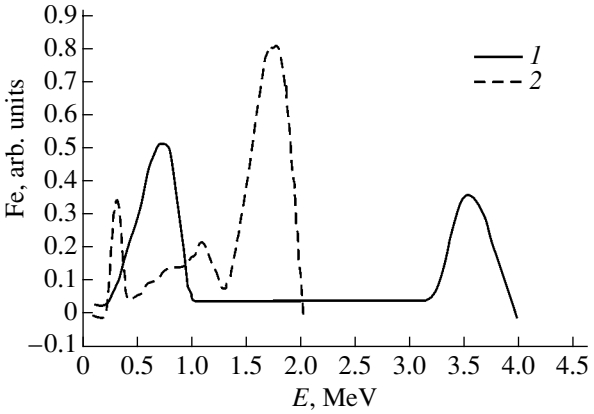


Fig. 5. Energy spectra of electrons in a UNI-10 accelerator: (*1*) short impulse and (*2*) long impulse.

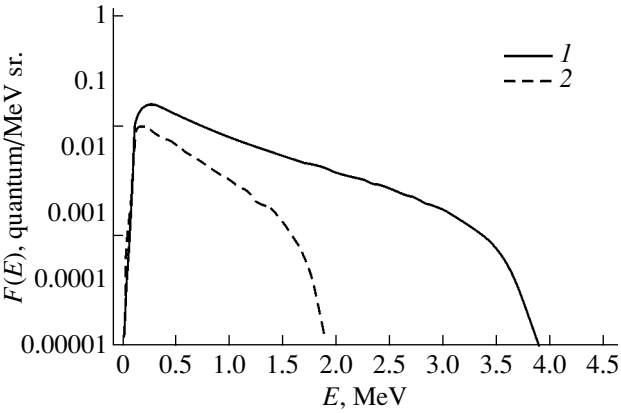


Fig. 6. Energy spectra of bremsstrahlung from a UNI-10 accelerator. The notation for the curves is identical to that in Fig. 5.

The energy spectra of electrons in radiation pulses of a UIN-10 accelerator for various modes of its operation are presented in Fig. 5 according to the results derived by using the converter-target transfer function [2, 18]. In the first case, a converter target from steel was employed to generate bremsstrahlung. A pulse was formed by using an explosive current interrupter, the effective pulse duration being $\tau_p \approx 80$ ns (short pulse). A converter target from tantalum was harnessed in the second case. The mode of a direct discharge of an inductive storage device to a vacuum diode was realized, the effective pulse duration being $\tau_p \approx 2$ μs in that case (long pulse).

The energy spectra F of bremsstrahlung for these modes of operation of a UIN-10 accelerator were determined for the forward direction (Fig. 6), the LUE and ANGLE1 codes being employed for, respectively, the short- and the long-pulse mode.

CONCLUSION

The proposed semiempirical methods for determining the energy spectra of bremsstrahlung from single-layer converter targets in pulsed electron accelerators make it possible to reduce the time of calculations by one (for the ANGLE1 code) or two (for the LUE code) orders of magnitude in relation to calculations by the Monte Carlo method [6]. Together with the methods developed in [1, 2] for measuring the angular and energy features of electrons in an accelerator radiation pulse, they make it possible to automate the derivation of information both in the diagnostics of the accelerator operation mode and in the dosimetry of the radiation tests of samples.

REFERENCES

1. N. G. Mordasov, Zh. Tekh. Fiz. **73** (2), 107 (2003) [Tech. Phys. **48**, 239 (2003)].
2. N. G. Mordasov, Vopr. At. Nauki Tekh., Ser.: Fiz. Radiats. Vozd. Radioélektron. Apparaturu, Nos. 3–4, 56 (2002).
3. V. P. Kovalev, *Secondary Radiation of Electron Accelerators* (Atomizdat, Moscow, 1979) [in Russian].
4. H. Tseng and R. Pratt, Phys. Rev. A **19**, 1525 (1979).
5. L. Schiff, Phys. Rev. **83**, 252 (1951).
6. G. V. Miloshevskiĭ, Inzh.-Fiz. Zh. **71**, 887 (2000).
7. T. Nakamura, M. Takemura, H. Hirayama, and T. Hyodo, J. Appl. Phys. **43**, 5189 (1972).
8. A. A. Astakhov and V. A. Balitskiĭ, in *Applied Nuclear Spectroscopy* (Énergoatomizdat, Leningrad, 1984), Vol. 13, pp. 118–122 [in Russian].
9. A. A. Astakhov and N. G. Mordasov, in *Proceedings of Conference on Radiation Resistance of Electronic Systems "Resistance-2003," Moscow, 2003* (Mosk. Inzh.-Fiz. Inst., Moscow, 2003), pp. 179–181.
10. A. Hanson, L. Lansl, E. Lyman, and M. Scott, Phys. Rev. **84**, 634 (1951).
11. H. Bethe, Phys. Rev. **89**, 1256 (1953).
12. *Beta- and Gamma-Ray Spectroscopy*, Ed. by K. Siegbahn (Interscience, New York, 1955; Inostrannaya Literatura, Moscow, 1959).
13. M. Berger and S. Seltzer, *Tables of Losses and Ranges of Electrons and Positrons* (NASA, Washington, 1964), SP-3012.
14. T. Tabata and R. Ito, Nucl. Instrum. Methods **127**, 429 (1975).
15. N. Hansen and S. Fultz, *Cross Section and Spectra for Negative Electron Bremsstrahlung, Lawrence Radiation,* Report No. UCRL-6099 (1960).
16. G. Moliere, Z. Naturforsch. A **3**, 78 (1948).
17. A. O'Dell, C. Sandifer, R. Knowlen, and W. George, Nucl. Instrum. Methods **61**, 340 (1968).
18. N. G. Mordasov, V. N. Ulimov, V. A. Bryksin, and V. D. Shiyan, in *Proceedings of the Conference on Radiation Resistance of Electronic Systems "Resistance-2003," Moscow, 2003* (Mosk. Inzh.-Fiz. Inst., Moscow, 2003), pp. 175–176.

Translated by A. Isaakyan

Technical Physics, Vol. 49, No. 9, 2004, pp. 1221–1223. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 116–119.

BRIEF COMMUNICATIONS

Thermal and Electrical Characteristics of Thermionic Converters under Combined Electrode Heating

Yu. I. Dudarev and B. I. Ermilov

Sukhumi Physicotechnical Institute, Academy of Sciences of Abkhazia, Sukhumi, 384914 Abkhaziya, Georgia

Received July 1, 2003

Abstract—An experimental procedure for studying the thermal and electric characteristics of a single-element thermionic converter under combined heating of the electrodes is proposed (direct heating by thermal radiation from an electric heater plus heating of the emitter by the reverse currents of the collector). The *I–V* characteristics of the converter are studied under various operating conditions and at different lengths of the reverse heating current pulse. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Heat sources used in direct thermoelectric converters are nuclear reactors, radioactive isotopes, and chemical and solar energy [1]. Since these sources are expensive and difficult to use, they are sometimes inappropriate in experimental studies of the physical, electrical, and thermal characteristics of the converters. Therefore, if possible, designers employ heat source simulators.

The main techniques of heating the emitter of a thermionic converter (TC) are heating by radiation, the application of reverse currents from the collector to the emitter, or a combined technique (radiation plus electron bombardment). The combined technique is widely used in testing laboratory TCs. All other things being equal, it significantly lowers the temperature of the heating element and, consequently, increases its service life.

DEVICE PROTOTYPE

As the test object, we used a model single-element TC shown in Fig. 1. It consists of an emitter, a collector, a bellows pressure-tight feedthrough, a heat exchanger, and a thermocouple well. The emitter is a cylindrical pipe made of a refractory material and covered by a thin tungsten layer from the outside. Three holes of different depth (arranged at an angle of 120° relative to each other) are drilled in the end face of the emitter in order to fix beryllia-sheathed tungsten–rhenium thermocouples of diameter d = 0.2 mm. The emitting surface of the collector is also made of tungsten. Chromel–Alumel thermocouples, which are inserted through the thermocouple well, are placed on the outside of the collector. The emitting area of the emitter is 10.4 cm^2, and the electrode gap is 0.25 mm. The heat exchanger is designed in the form of a gas–water cooling jacket. The gaseous gap is 0.4 mm wide.

The emitter is heated by the main rod-type electric heater (MEH) inserted into the cavity. Desired thermal conditions are provided by Nichrome-wire electric heaters mounted on the heat exchanger and bellows pressure-tight feedthrough.

EXPERIMENTAL TECHNIQUE AND RESULTS

In the experiments, the electrodes of the single-element TC were subjected to combined heating, which included (i) direct heating of the emitter by thermal radiation with a resistance heater placed inside the emitter and (ii) heating by the reverse currents from the collector, which is heated by an external heater ("reverse heating"). In the latter case, the emitter is heated by the kinetic energy of the electron flow from the arc-discharge plasma.

The electrical performance of the TC under the combined heating of the emitter, as well comparison of the characteristics taken under direct and combined heatings, was studied with the equipment intended for TC heating and diagnostics by the reverse current technique [2]. Preparatory to testing the TC, we evacuated

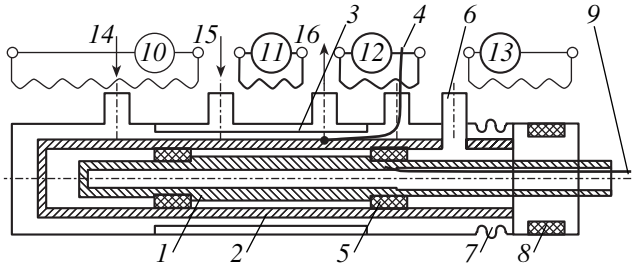


Fig. 1. Schematic of the device: (*1*) emitter, (*2*) collector, (*3*) water jacket, (*4*) collector thermocouple, (*5*) spacers, (*6*) electrode gap evacuation outlet, (*7*) bellows, (*8*) insulator, (*9*) emitter thermocouple, (*10–13*) heaters, (*14*) gas inlet, (*15*) water inlet, and (*16*) water outlet.

1063-7842/04/4909-1221$26.00 © 2004 MAIK "Nauka/Interperiodica"

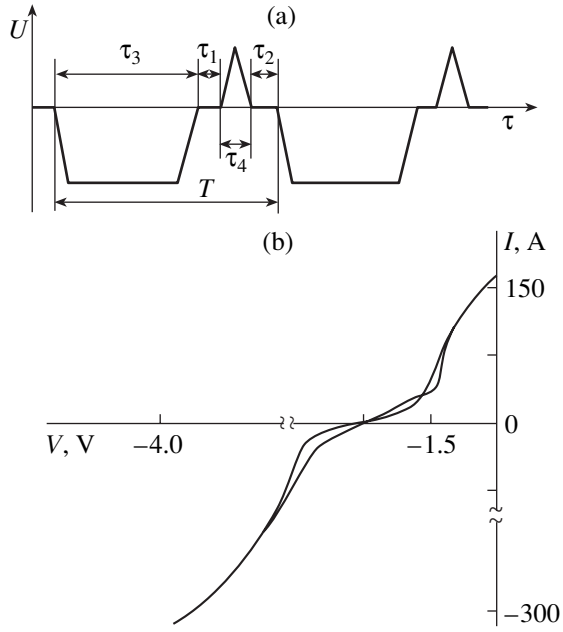


Fig. 2. (a) Pulse waveform and (b) *I–V* characteristic of the single-element TC under the reverse heating. τ_3 is the heating pulse length, τ_4 is the measuring pulse length, τ_1 and τ_2 are the time delays, and *T* is pulse repetition period.

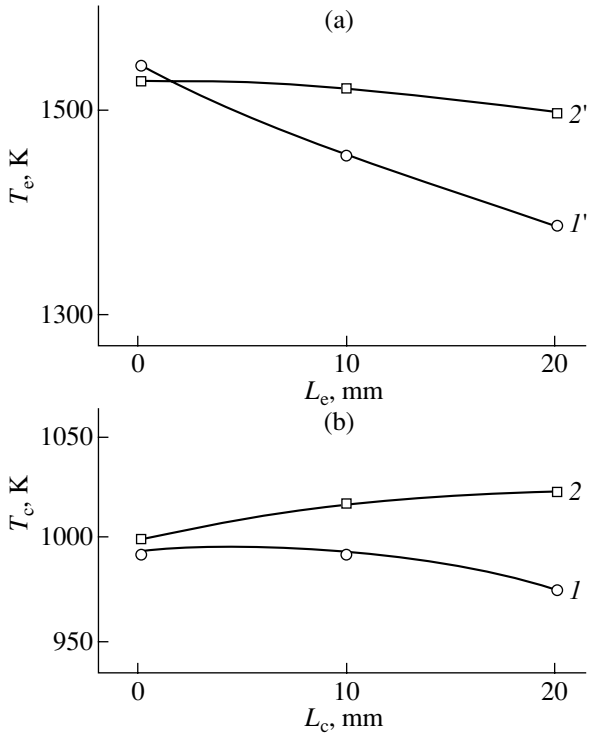


Fig. 3. Temperature distribution along the (a) emitter and (b) collector under the reverse heating: (*1*, *1*') the powers of the external heater sections are the same and (*2*, *2*') the power of the right-hand section is higher.

the vacuum–cesium system and degassed the electrodes using the rod-type MEH. The vacuum–cesium system was evacuated at a channel temperature of 700–750 K until a pressure of $P_1 = 1 \times 10^{-4}$ Pa (dynamic vacuum) was established. Under these conditions, the pressure in the gaseous gap was $P_1 = (4–6) \times 10^{-3}$ Pa.

The electrodes were degassed in two steps. First, the collector was degassed as follows: its temperature was increased with the help of the MEH to 1100 K at a rate limited by the pressure in the electrode gap, no more than 1×10^{-4} Pa, and in the gaseous gap, $(5–6) \times 10^{-3}$ Pa. These conditions were maintained until the pressure in the electrode gap was kept at $P_2 = 1 \times 10^{-4}$ Pa for 1 h. After the collector had been degassed, the gaseous cavity of the TC was filled with helium to a pressure of 100–260 Pa with the MEH switched off. At the second step, the emitter was degassed with the MEH at a temperature of 1900 K and a pressure in the electrode gap of $\sim 1 \times 10^{-4}$ Pa. These conditions were maintained until the pressure $P_2 \approx 1 \times 10^{-4}$ Pa was kept for 1 h. The reference *I–V* characteristics (direct heating) were taken with the electrodes MEH-heated to T_e = 1400–1900 K (emitter) and T_c = 950–1100 K (collector) and the temperature of the working medium (cesium) varying between 520 and 620 K.

Under the combined heating of the device, the water jacket of the heat exchanger was connected to the air delivery system, which made air circulate through the jacket under a pressure of 110 to 150 kPa and helium was delivered to the gaseous gap at a pressure of $(2–5) \times 10^{-3}$ Pa. The external heater was used to raise the collector temperature to 770–870 K. A dc accelerating voltage of –12 V was applied to the lead terminals of the device, with the emitter terminal charged positively.

The experimental heating equipment provided discharge current stabilization (the current could be varied between 0 and 300 A); therefore, as the electrodes were heated, the accelerating voltage was decreased. When the temperature of the cesium thermostat reached 570–620 K, the electrode gap was filled with the cesium vapor. Such a high temperature of the thermostat excludes stray breakdowns between neighboring elements that are separated by a distance greater than the electrode spacing.

To gain information on TC performance from the *I–V* characteristics, reverse heating was accomplished in the pulsed mode. The *I–V* characteristics were taken during intervals between the pulses.

Figure 2 shows the time waveform of the heating and diagnostic pulses and also the reference *I–V* curve of the single-element TC under reverse heating. Under reverse heating, the *I–V* curve exhibits hysteresis, which is typical of an arc discharge. This fact evidences that an arc discharge is initiated and quenched between the electrodes during the heating pulses. When the electrodes operate under different thermal conditions (direct heating, reverse heating, and their combination), the temperature profiles of both vary. If the temperature

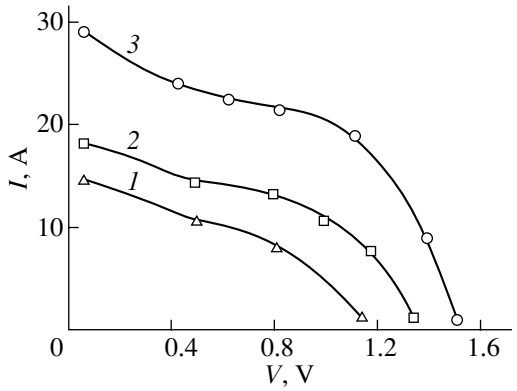


Fig. 4. *I–V* characteristics of the TC under different heating conditions at T_e = 1720 K, T_c = 1030 K, and P_{cs} = 0.9 GPa. (*1*) Direct heating at W_1 = 834 W, (*2*) combined heating at W_1 = 670 W and W_2 = 212 W, and (*3*) combined heating at W_1 = 350 W and W_2 = 528 W.

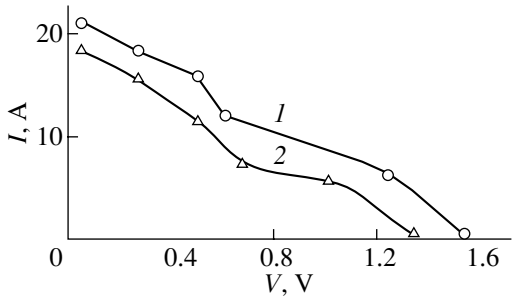


Fig. 5. *I–V* characteristics of the TC for different heating pulse lengths τ_3 at P_{cs} = 1.9 GPa. (*1*) τ_3 = 46 ms, τ_1 = 18 ms, τ_4 = 10 ms, and T = 104 ms and (*2*) τ_3 = 18 ms, τ_1 = 40 ms, τ_2 = 26 ms, τ_4 = 10 ms, and T = 100 ms. Combined heating at (*1*) W_1 = 865 W, W_2 = 278 W, T_e = 1820 K, and T_c = 1110 K and (*2*) W_1 = 997 W, W_2 = 264 W, T_e = 1820 K, and T_c = 1110 K.

profile of the collector is changed using the tapped heaters of the heat exchanger and the electric current through the TC is stabilized, the density of the electric current from the collector is redistributed under the reverse heating, which changes the emitter temperature in the corresponding segment.

Curve *1* in Fig. 3 shows the temperature distribution along the collector at a cesium pressure P_{cs} = 4.5 GPa when the powers of both sections of the heaters of the heat exchanger are equal. When the power of the right-hand section of the heater increases, causing a corresponding increase in the collector temperature in this region by 40 K (curve *2*), the change in the emitter temperature in the related segment is ≈100 K.

The experiments confirmed the possibility of controllably initiating an arc discharge at a desired site under reverse heating by creating a nonisothermal profile of the collector. This allows emitter temperature profiling under combined and reverse heating and, in particular, makes it possible to simulate heat generation at the electrodes of the electric-power-generating channel in out-of-pile tests.

The electrical performance of the TC was studied under different conditions of electrode heating. Figure 4 shows the *I–V* characteristics taken under direct and combined heating. The electrode temperatures were $T_e \approx 1720$ K and $T_c \approx 1030$ K, and the cesium pressure was P_{cs} = 0.9 GPa. It is seen that, as the reverse heating component of the heat balance grows under combined heating, the *I–V* curve shifts toward higher voltages approximately by 0.45 V and the converter's output increases significantly (by a factor of 2 to 3). These findings probably indicate that the discharge with distributed thermionic emission from the emitter changes to the cathode-spot discharge [3] although the electrodes are short (L = 20 mm). Similar results were observed when the characteristics were taken under the same thermal conditions but at higher cesium vapor pressures.

We also studied the effect of the heating pulse length (τ_3) on the converter's output parameters. Figure 5 shows the *I–V* characteristics of the converter that are taken at different heating pulse lengths. As follows from Fig. 5, the output parameters measured at τ_3 = 46 ms are almost 1.5 times higher than those measured at τ_3 = 18 ms. Specifically, the emf is higher by approximately 0.2 V in the former case.

CONCLUSIONS

(1) A thermal test procedure for single-element cylindrical TCs under direct and combined electrode heatings is presented.

(2) Analysis of the converter's electrical performance indicates the possibility of controllably initiating an arc discharge under the reverse heating by producing a nonisothermal temperature profile over the collector.

(3) It is shown that, under the combined heating, the output parameters of the converter are considerably improved as the reverse heating component of the heat balance grows.

(4) Under the combined heating of electrodes of such geometry, the heating pulse length τ_3 = 46 ms is found to be optimal.

REFERENCES

1. F. G. Baksht, G. A. Dyuzhev, A. M. Martsinovskiĭ, *et al.*, in *Thermionic Converters and Low-Temperature Plasma*, Ed. by B. Ya. Moĭzhes and G. E. Pikus (Nauka, Moscow, 1973) [in Russian].
2. B. N. Igumnov and V. P. Evilin, Élektrichestvo, No. 4, 80 (1976).
3. G. A. Dyuzhev, V. B. Kaplan, B. Ya. Moĭzhes, *et al.*, Zh. Tekh. Fiz. **41**, 453 (1971) [Sov. Phys. Tech. Phys. **16**, 584 (1971)].

Translated by A. Khzmalyan

Technical Physics, Vol. 49, No. 9, 2004, pp. 1224–1226. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 120–121.

BRIEF COMMUNICATIONS

Beam Instability of Afterglow Plasma in Inert Gases

A. S. Mustafaev

St. Petersburg State Mining Institute (Technical University),

St. Petersburg, 199026 Russia

Received December 30, 2003

Abstract—The paper considers the conditions under which the development of beam instabilities of the afterglow plasma in inert gases is possible. The validity of assumptions has been experimentally confirmed. © *2004 MAIK "Nauka/Interperiodica".*

A system consisting of a plasma and a beam of charged particles passing through it is known to become unstable under certain conditions [1, 2]. It was shown [3–5] that such instability arises in a plasma even in the presence of weak anisotropy of the electron distribution function (EDF) with respect to velocities. Traditionally, investigations devoted to the EDF of electrons in decaying plasmas of inert gases [6–8] did not analyze the process of instability development and its influence on the form of the EDF.

This paper considers for the first time the possibility of the development of beam instability of a decaying afterglow plasma in inert gases, where fast electrons appear as a result of physical and chemical processes.

Let us consider an afterglow plasma in an inert gas, where electrons are produced as a result of the chemical ionization reactions

$$A^* + A^* \begin{cases} A^+ + A + \mathbf{e} \\ A_2^+ + \mathbf{e} \end{cases} \quad (1)$$

and collisions of the second kind between excited atoms and electrons

$$A^* + e \longrightarrow A + \mathbf{e}. \quad (2)$$

Here, A and A^* are atoms in the ground and excited states; A^+, A_2^+ are atomic and molecular ions; and **e** (e) are fast (slow) electrons. The energies of the fast electrons produced during the above physical and chemical processes in inert gases, ε_1 (reaction (1)) and ε_2 (reaction (2)), are listed in the table.

The ambipolar electric field in plasma is determined by the averaged electron energy and weakly influences the diffusion of fast electrons. Under these conditions, the preferential motion of fast electrons toward plasma volume boundaries introduces anisotropy in the EDF, although the fast electron "source" itself is isotropic. The degree of anisotropy of the EDF is determined by the gas pressure, the plasma volume, and the electron density. When the "critical" electron density is reached, plasma–beam instability develops. This instability leads to energy relaxation of the EDF and to smearing of the EDF peaks near the energies ε_1 and ε_2 [9].

Let us analyze the results of measurements of the EDF in an afterglow plasma in helium. The experiments were carried out in a discharge tube with a diameter of 2.7 cm at a gas pressure of 1 Torr [7]. The plasma was produced by current pulses with an amplitude from 0.1 to 1.0 A, a duration of 10 μs, and a repetition frequency of 2 kHz. The EDF with respect to energies was measured by the method of probe current modulation with a time resolution of 10 μs in 200 μs after termination of the current pulse. The densities of electrons (n) and metastable helium atoms (M) in the 2^3S_1 state were simultaneously recorded.

Figure 1 shows the EDF measured in the afterglow plasma in helium for various values of the electron density n. As the electron density increases, the Langmuir wavelength λ decreases and the development of plasma–beam instability is facilitated [2, 9]. It is seen

Table

Gas	ε_1, eV	ε_2, eV
Helium	14.4	19.8
Neon	11.0	16.7
Argon	7.3	11.6
Krypton	6.0	10.0
Xenon	4.3	8.4

1063-7842/04/4909-1224$26.00

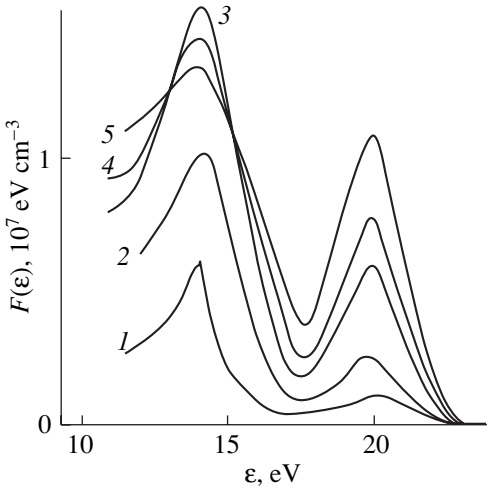


Fig. 1. Electron distribution function $F(\varepsilon)$ with respect to energies in afterglow plasma in helium for $n = 4.5 \times 10^{10}$ cm^{-3} (*1*), 10^{11} (*2*), 2×10^{11} (*3*), 3×10^{11} (*4*), and 4×10^{11} (*5*).

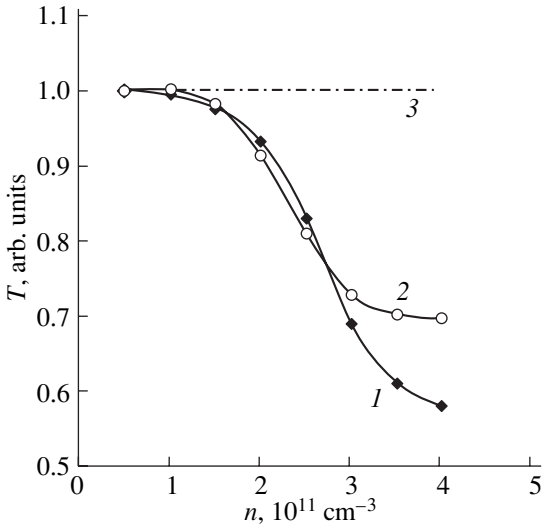


Fig. 2. The typical lifetime of fast electrons T as a function of n for electrons with the energy (*1*) $\varepsilon_1 = 14.4$ eV and (*2*) $\varepsilon_2 = 19.8$ eV; (*3*) calculation with neglect of the plasma wave excitation.

that the groups of fast electrons with energies $\varepsilon_1 = 14.4$ eV and $\varepsilon_2 = 19.8$ eV are actually present in the energy spectrum. The dispersion of EDF for the first electron group ($\Delta\varepsilon_1$) is appreciably greater than that for the second group ($\Delta\varepsilon_2$). This fact is apparently associated with that the energy ε_1 is lower than ε_2 and, correspondingly, the excitation cross section of helium atoms by electrons of the latter group is smaller. This, in turn, leads to a higher diffusion length $L_1 > L_2$ and, hence, to a higher lifetime for the electrons of this group.

For a comparative analysis of the transformation of the shape of the distribution function depending on the electron density, let us consider the experimental dependence of the typical lifetime T of fast electrons shown in Fig. 2 in relative units. The lifetimes T are proportional to the ratios of amplitudes of the EDF peaks at energies of 14.4 and 19.8 eV to the corresponding rates of generation of the electrons of these groups. For the conditions considered (excitation of waves in a plasma is neglected), this time should be independent of n (dash line). The rate constants of the corresponding processes and the rates of generation of fast electrons were determined using the measured values of densities n and M [8].

Analysis of the experimental results shows that appreciable transformation of the shape of the EDF in decaying helium plasma occurs even for $n \geq 3 \times 10^{11}$ cm^{-3}. This result is in good agreement with a threshold criterion for the development of plasma–beam instability and the following energy relaxation of the EDF. This criterion was found for the plasma of low-voltage beam discharge in inert gases [9]. It should be noted that the observed changes in the form of the EDF with increasing density n are substantially greater than the experimental errors, which did not exceed 5% for relative measurements.

The condition of beam instability development in decaying afterglow plasma in inert gases can conveniently be expressed as

$$n \geq 1.7 \times 10^8 \varepsilon \sigma^2(\varepsilon) N_A^2, \qquad (3)$$

where ε is the energy of fast electrons [eV], n is the density of thermal electrons [cm^{-3}], N_A is the density of normal gas atoms [cm^{-3}], and $\sigma(\varepsilon)$ is the cross section of elastic electron scattering on atoms [cm^2].

Substitution of the corresponding values into Eq. (3) leads to the experimental value of n corresponding to the onset of energy relaxation of the EDF in a decaying plasma.

To summarize, the analysis performed leads to the conclusion that there is beam instability development is possible in an afterglow plasma in inert gases under certain conditions. The beam instability changes the shape of the EDF with respect to energies and, hence, the properties of inert gas plasma. This circumstance should be taken into account in technologies using plasmas with fast electrons.

ACKNOWLEDGMENTS

The author is grateful to Prof. G. A. Dyuzhev for useful critical remarks.

REFERENCES

1. A. A. Vedenov, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1963), Vol. 3, pp. 203–244.
2. A. A. Ivanov, *Physics of Highly Nonequilibrium Plasma* (Atomizdat, Moscow, 1977) [in Russian].
3. A. S. Mustafaev and A. P. Mezentsev, J. Phys. D **19** (5), L69 (1986).
4. F. G. Baksht, V. F. Lapshin, and A. S. Mustafaev, J. Phys. D **28**, 689 (1995).
5. F. G. Baksht, V. F. Lapshin, and A. S. Mustafaev, J. Phys. D **28**, 694 (1995).
6. V. I. Demidov, N. B. Kolokolov, and O. G. Toronov, Fiz. Plazmy **12**, 702 (1986) [Sov. J. Plasma Phys. **12**, 402 (1986)].
7. V. I. Demidov and N. B. Kolokolov, Phys. Lett. A **89A**, 397 (1982).
8. N. B. Kolokolov, in *Chemistry of Plasma*, Ed. by B. M. Smirnov (Énergoatomizdat, Moscow, 1985), No. 12, pp. 56–66 [in Russian].
9. A. S. Mustafaev, Zh. Tekh. Fiz. **71** (4), 111 (2001) [Tech. Phys. **46**, 472 (2001)].

Translated by M. Astrov

Technical Physics, Vol. 49, No. 9, 2004, pp. 1227–1231. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 122–126.

BRIEF COMMUNICATIONS

Spatial Macroscopic Structure of a Pulsed Microsecond Discharge in Barrier–Metal Gaps with a Nonuniform Electric Field Distribution

S. N. Buranov, V. V. Gorokhov, V. I. Karelin, and A. A. Tren'kin

All-Russia Research Institute of Experimental Physics, Russian Federal Nuclear Center, Sarov, Nizhegorodskaya oblast, 607190 Russia

e-mail: karelin@ntc.vniief.ru

Received December 1, 2003

Abstract—Electric and spatial characteristics of a pulse-periodic microsecond barrier discharge are investigated in different geometries—triangular prism, plate, and corrugated electrode—that are in contact with a dielectric plate and form a dihedral angle with it. It is established that, in the space of the dihedral angle, the regions of discharge represent alternating cylindrical layers with the axes lying on the contact line. The first conducting layer is formed at some distance from the contact edge of the electrode. The number of layers and their localization are determined by the angle formed between an electrode and the dielectric plate. A physical model explaining the main features of the structure formation is proposed. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Investigation of barrier discharge is of special interest due to its wide applications: ozone synthesis, destruction of detrimental impurities in air, production of excimers, and organic and inorganic synthesis [1–3]. Recently, a new line of research in medicine —ozone therapy—is being intensively developed; the technological equipment in this field requires the creation of special instrumentation—medical ozonizers [4].

An important requirement specific to medical ozonizers is the stability of ozone production in a wide range of output concentrations. In this case, a challenge is presented by accurate stable synthesis of O_3 of low concentrations. To solve this problem, it was necessary to pass to microsecond pulses for feeding the discharge and the breakdown stabilization of a gaseous gap in each pulse [5]. Investigations indicate [5, 6] that, when feeding microsecond pulses, it is very difficult to satisfy these requirements in gaps with a uniform field geometry, since high voltages, high accuracy of the preparation of electrode systems, and the introduction of initiating electrodes are needed. These circumstances dictate the passage to discharge chambers with nonuniform field geometry [6]. At the same time, it should be noted that, even now, the number of studies devoted to investigating the barrier discharge excited by microsecond pulses even with a uniform field distribution in the gaps is rather limited, and the discharge in the cells with a nonuniform field distribution has hardly been investigated at all [5–7].

The aim of this study is to investigate the electric and spatial characteristics of the discharge excited by microsecond pulses in barrier–metal gaps with nonuniform field geometry.

EXPERIMENTAL GEAR

The discharge was studied in air (at atmospheric pressure) using a stand (Fig. 1) involving a power supply (*PS*), electrode system (*ES*), and diagnostic instrumentation. As a power supply, we used a transistor oscillator of alternating high-voltage pulses that was developed to operate with a nonlinear resistive–capacitive load. The oscillator was designed according to the scheme of a resonance inverter with subsequent magnetic front contraction and enhancement of the pulse amplitude by a high-voltage transformer. Each of the pulses consisted of two half-waves of opposite polarity. The parameters of the oscillator are as follows: the energy in a pulse is 16 mJ, the open-circuit voltage is 22 kV, the pulse rise time is 700 ns, the duration of the first half-wave at the half-height is 3.5 μs, and the repetition rate of pulses range from a single pulse to 25 Hz.

The electrode system mounted in a chamber (*C*) supported by dielectric insulators (*D.s*) consists of plane grid *1* and profiled *2* electrodes separated by a dielectric barrier *3*. Electrode *1* is made of a nickel 150×150 μm mesh grid (the diameter of wire is 80 μm). The geometric transparency of the grid is 5%. The grid is welded to the end of a metal ring (outer diameter, 50 mm) and tightly pressed to the barrier. As the barrier, we used a quartz plate of 1 mm in thickness and 60 mm in diameter. We used several configurations of electrode *2* that provided a nonuniform field distribution in the gap: triangle prisms, a thin (0.3 mm) rectangular plate, and corrugated electrode. In all the cases,

1063-7842/04/4909-1227$26.00

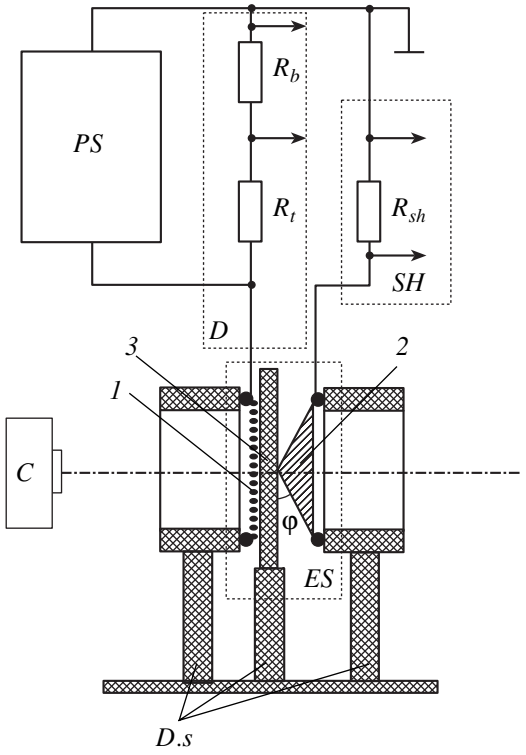


Fig. 1. Scheme of the laboratory stand.

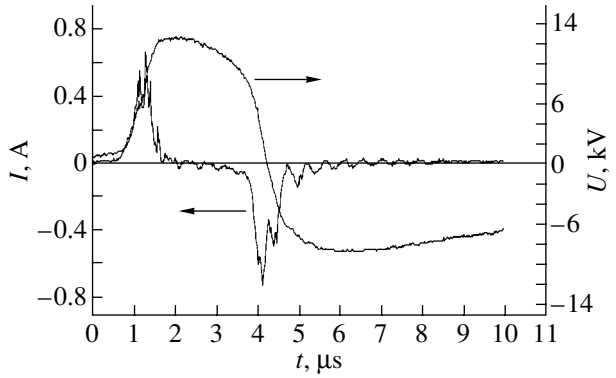


Fig. 2. Oscilliograms of voltage *U* and current *I* in the discharge (bared prism electrode, φ = 15°).

the electrode edges were in contact with the plate surface, forming a dihedral angle φ. The characteristic sizes of electrodes *2* did not exceed the diameter of the grid electrode. The range of the angles under investigation was between 15° and 60°.

The voltage *U* applied across the discharge gap and the discharge current *I* were recorded by using a Tektronix TDS754C oscilloscope having a pass band of 500 MHz. The time resolution of the resistive voltage divider *D* and the low-inductive coaxial shunt *SH* was no worse than 10 and 1 ns, respectively.

The discharge glow was photographed through the grid electrode and from the side in the direction of the contacting electrode–barrier line. The generator operated at a frequency of 25 Hz with an exposure time of 0.5–1 min.

To visualize the current distribution, the surface of electrode *2* was coated with a thin layer of soot (~1 μm). Under the action of the discharge in the ranges of current, the soot evaporated, thereby leaving the electrode surface bare. During evaporation, a part of the soot was deposited on the quartz plate. As a result, a photograph was developed that showed the current distribution on the plate surface. To obtain a clear picture of the distribution, the electrodes were treated for 1 min at the generator frequency of 25 Hz.

EXPERIMENTAL RESULTS

Our experiments showed stable discharge formation in the barrier–metal gaps with a nonuniform field distribution in each pulse, starting from single pulses. After applying a voltage pulse to the discharge gap, the discharge current appeared on both the leading and trailing edges of the first half-wave of a voltage pulse. The breakdown voltage at the leading edge of the pulse of both polarities varied from pulse to pulse in a range of 4–6 kV.

The electric characteristics were mainly investigated in the experiments with the prism electrodes. Figure 2 shows typical oscillograms of voltage *U* and discharge current *I*. The smooth rise of the current oscillogram on the initial segment is caused by the charging of the gap capacitance. When the voltage attains its breakdown value, a sharp rise in current occurs. The current achieves its amplitude value of 0.6–2 A in 5–25 ns, after which it drops to zero over 400–500 ns. At the trailing edge of the first half-wave of the voltage pulse, the current smoothly increases, attains its amplitude value, and then smoothly drops. The duration of the current pulse equals 800 ns. This is explained by recharging of the barrier capacitance.

In our experiments, a detailed investigation of the dependence of the electric characteristics of the discharge on the angle φ was not performed. However, the analysis of the results obtained indicates that the above dependence is absent.

It should be noted that the soot deposited on the electrode surface has a considerable effect on the electric characteristics of the discharge: the breakdown voltages decreased (3.2–3.7 kV) and the discharge current increased more smoothly, without a pronounced amplitude step.

The discharge was photographed when operating with electrode *2* made in the form of a plate. The discharge photographs obtained through the grid electrode and from the side are presented in Figs. 3a and 3b, respectively. The crosshatched region in Fig. 3a corresponds to the contact line and, in Fig 3b, the position of the electrode. In Fig. 3a, one can observe alternating regions of the discharge glow in the form of strips ori-

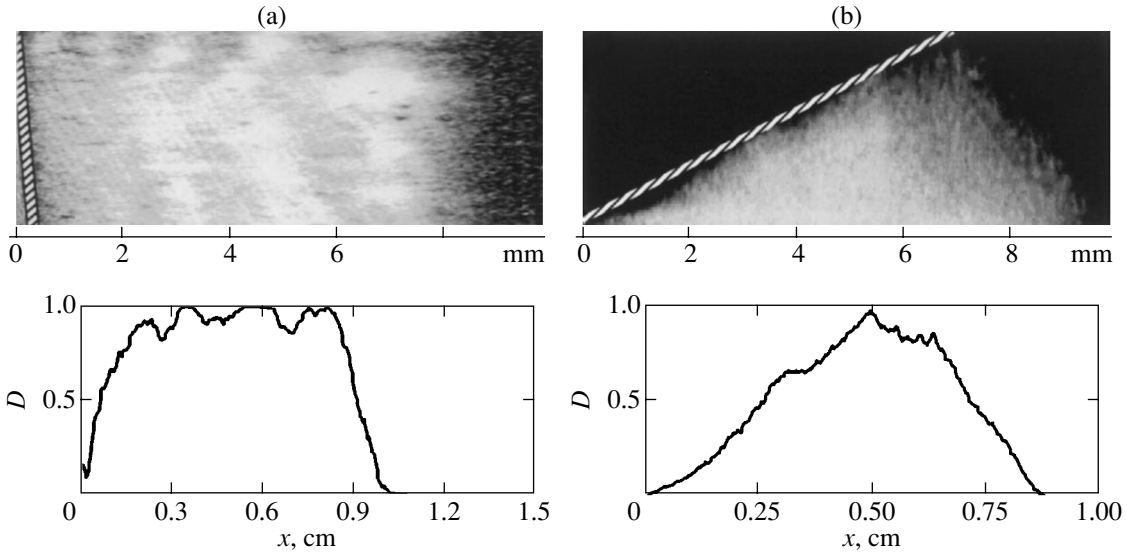


Fig. 3. Photographs of the discharge and corresponding densitograms (plate, φ = 15°).

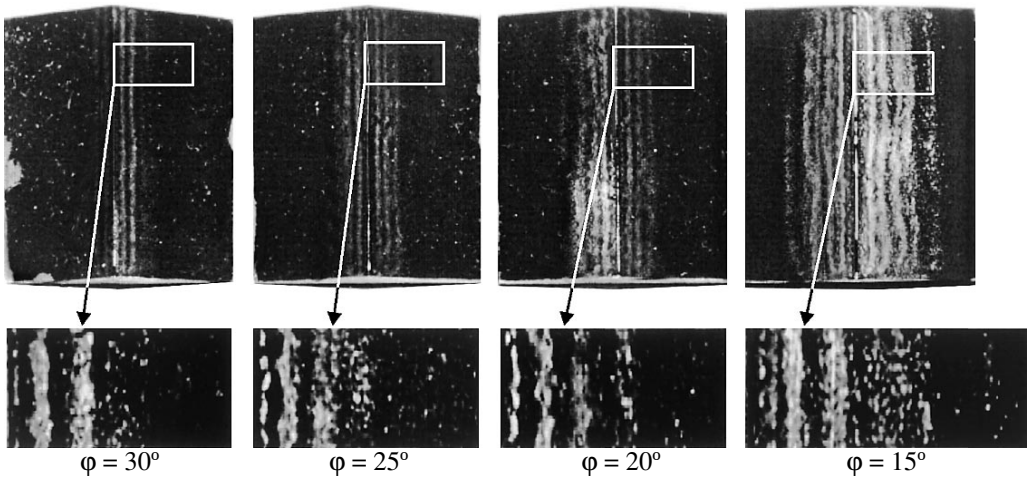


Fig. 4. Electrodes after treatment in the discharge.

ented in the direction of the contact line. The bending of the strips is caused by the deformation of the electrode surface.

When photographing from the side (along the contact electrode–dielectric line), the discharge glow averaged over the entire length of the interelectrode gap is recorded. Therefore, the structure of the discharge in this projection (Fig. 3b) is less pronounced. Nevertheless, one can detect the regions of maximum glowing that are close in form to the circular arcs with the centers at the contact point. Comparison of Figs. 3a and 3b indicates that, for both projections, the structural features of the discharge are similar.

The photographs of the soot-coated prism electrodes, after their work surfaces were treated in the discharge, are shown in Fig. 4. As is seen, the erosion region represents a set of strips virtually parallel and symmetric to the contact line. However, it should be noted that the structure recorded in our experiments is macroscopic with respect to the barrier-discharge microstructure [1, 2, 6], which was not investigated in this study. The erosion action of the discharge is proportional to the current density; therefore, the light strips correspond to the regions with higher current density, whereas the dark strips correspond to the regions where the discharge current is small or even absent. Figure 5 shows a photograph of the current distribution over the dielectric surface. Here, the region of the discharge also alternate with the regions similar in structure, where the discharge is absent. Taking into account the discharge structure photographed from the side (along the contact line (Fig. 3b)), one can conclude that, in the bulk of the discharge gap, the current

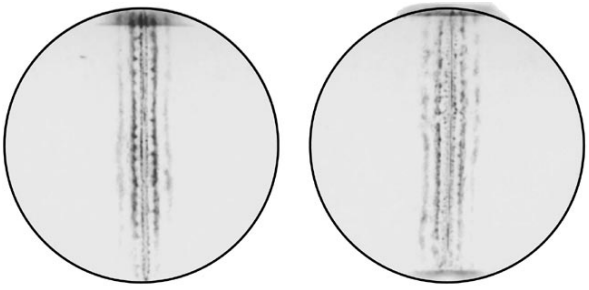


Fig. 5. Dielectric plate surface after treatment in the discharge.

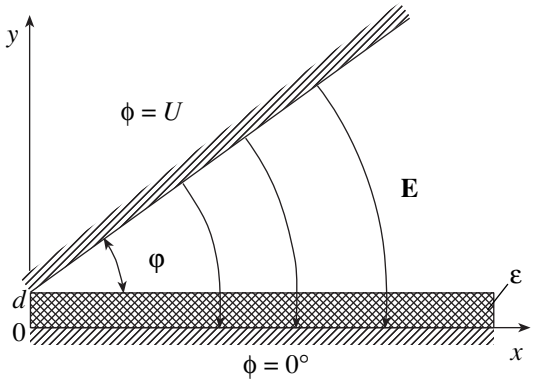


Fig. 6. Model of the electrode system (*d* is the barrier thickness and ε is the permittivity).

regions of the discharge are the cylindrical layers with axes coinciding with the contact line.

The first erosion strip on the surface of electrode *2* (Fig. 4) and the first light strip on the photographs (Fig. 3), which correspond to the position of the first current layer, are located at some distance from the contact line. Consequently, the first conducting region is formed within the gas gap (at some distance from the contact line) rather than from the sharp edge of electrode *2*, which is in contact with the barrier. The location and the number of the current layers depend on the angle φ. With increasing angle φ, the layers are shifted to the line contact and their number decreases. Starting from an angle of φ_k ($50° < \varphi_k < 60°$), only one layer distant from the electrode edge is left and an incomplete sliding discharge is observed in the gap (along the dielectric surface).

Distances between the contact line and *i*th strip

φ, deg	x_i, cm				
	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
15	0.14	0.28	0.5	0.7	0.9
20	0.09	0.23	0.42	0.63	0.92
25	0.08	0.2	0.37	0.55	0.85
30	0.05	0.17	0.37	0.55	–

The averaged distances x_i ($i = 1, \ldots, 5$) between the electrode edge and *i*th current layer, which were obtained by image processing, are presented in the table. The dependence of the distances x_1, x_2, x_3, and x_4 on φ can be approximately described by a hyperbolic law.

A similar structure was also observed in the experiments with corrugated electrode *2* (see Fig. 1).

FORMATION OF DISCHARGE STRUCTURE IN BARRIER–METAL GAPS WITH NONUNIFORM FIELD DISTRIBUTION

To analyze the process of structural formation, we used a model of an electrode system formed by infinitely large metal and dielectric plates (Fig. 6). In such a model, at small angles φ ($\varphi < 1$), the electric field distribution in the range of $d < y < d + \varphi x$ can be approximated by the expression

$$E(x) = \frac{U}{\varphi x + \frac{d}{\varepsilon}}.$$

For each particular *x*, this expression describes the field inside a plane capacitor with a dielectric plate coating one of the capacitor plates. The numeric calculation of the field in the electrode system gives close values of electric field strength and indicates that the normal field component significantly exceeds the tangential field component. In this connection, the field lines we will represent by segments parallel to the 0*y* axis.

At the initial stage of the discharge (in the absence of a space discharge and a charge on the dielectric surface), the motion of electrons in the gap is drift. The electrons move along electric field lines. The breakdown of the gap is most probable in a region where an electron produces the maximal number of ionization events. As a parameter characterizing the ionization rate at the stage of avalanchelike generation of electrons, we will use the integral multiplication factor [8]

$$\chi(l) = \int_l \alpha(E)dl, \qquad (1)$$

where

$$\alpha(E) = pA\exp\left(-\frac{Bp}{E}\right)$$

is the Townsend effective ionization coefficient [9] (*p* is the gas pressure and $A = 15$ and $B = 365$ are constants) and the integration is performed along the field lines.

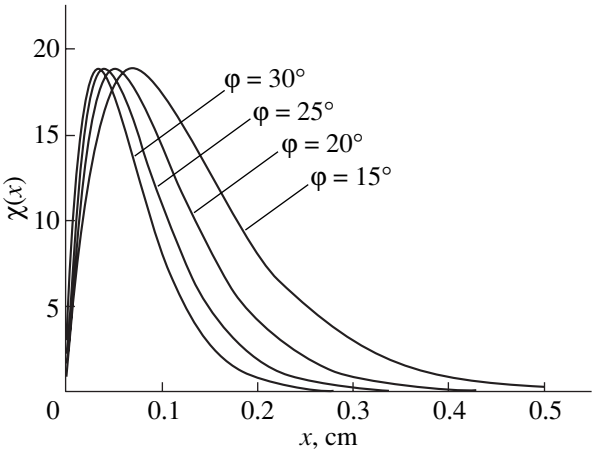


Fig. 7. Distribution of the integral multiplication factor of electrons $\chi(x)$ at the time instant of the gap breakdown.

Integrating (1), we obtain

$$\chi(x) = \varphi x p A \exp\left[-\frac{Bp}{U_1}\left(\varphi x + \frac{d}{\varepsilon}\right)\right].$$

With allowance for the experimental data, the graph of the dependence $\chi(x)$ at the time instant of the gap breakdown ($U_1 = 5$ kV) for different angles is presented in Fig. 7.

The position of the first current layer is determined by the region of the gap x_1 where the function $\chi(x)$ reaches its maximum:

$$x_1 = \frac{U_1}{p\varphi B}. \quad (2)$$

It is seen that function $\chi(x)$ reaches its maximum at some distance x_1 from the contact line; therefore, the discharge is initiated in the region that is distant from the contact line between the metal plate and the dielectric plate. The maximum value of function $\chi_{max} = \chi(x_1)$ is independent of angle φ that testifies to the existence of a threshold value of χ_n, which is the criterion of breakdown in this case [8].

A charge transferred by the conduction current on the dielectric surface is spread over the plate, thereby suppressing the electric field near the layer being formed and hampering the development of conduction in the adjoining regions. A further increase in voltage U across the electrodes leads to the formation of a new conduction region in the bulk of the gas gap (at a distance x_2 from the contact line), where the effect of the surface charge is insignificant. The process is repeated as long as the condition $U \geq x_n \varphi E_{bd}$ is fulfilled (E_{bd} is the intensity of the gas breakdown and x_n determines the position of the next conduction region).

The values of the quantities x_1 obtained from Eq. (2) are less than those obtained from the experiment. This can be explained by the fact that the actual electric fields near the contact line of the electrode edge and dielectric plate exceed the model ones.

CONCLUSIONS

Thus, we have investigated the electric and spatial characteristics of the pulse-periodic discharge excited in barrier–metal gaps with nonuniform field geometry.

It was found that the discharge arises at some distance from the contact line between the electrode and dielectric plate and forms a spatial structure consisting of cylindrical layers parallel to the contact line.

We proposed a physical model of the discharge formation. The model explains the main experimental data: the development of the discharge at some distance from the contact electrode–dielectric line, appearance of the spatial structure, and hyperbolic character of the x_1 vs. φ dependence.

REFERENCES

1. Yu. V. Filippov, V. A. Voblikova, and V. I. Panteleev, *Electrical Synthesis of Ozone* (Nauka, Moscow, 1987) [in Russian].
2. V. G. Samoĭlovich, V. Ch. Gibalov, and K. V. Kozlov, *Physical Chemistry of Barrier Discharge* (Izd. Mosk. Gos. Univ., Moscow, 1989) [in Russian].
3. S. P. Bugaev, V. A. Kuvshinov, N. S. Sochugov, *et al.*, Pis'ma Zh. Tekh. Fiz. **22** (17), 49 (1996) [Tech. Phys. Lett. **22**, 711 (1996)].
4. S. P. Peretyagin, in *Proceedings of the 2nd All-Russia Scientific–Practical Conference "Ozone in Biology and Medicine," Nizhni Novgorod, 1995*, pp. 25–26.
5. S. N. Buranov, V. V. Gorokhov, V. I. Karelin, *et al.*, in *Proceedings of the 2nd All-Russia Scientific–Practical Conference "Ozone in Biology and Medicine," Nizhni Novgorod, 1995*.
6. S. N. Buranov, V. V. Gorokhov, V. I. Karelin, *et al.*, in *Proceedings of the 12th IEEE International Pulsed Power Conference, Monterey, 1999*, pp. 1421–1424.
7. V. V. Shalimov, S. A. Dauengauer, E. B. Bespalova, *et al.*, Zh. Tekh. Fiz. **61** (6), 165 (1991) [Sov. Phys. Tech. Phys. **36**, 685 (1991)].
8. D. V. Razevig and M. V. Sokolova, *Calculation of Initial and Discharge Voltages of Gas Gaps* (Énergiya, Moscow, 1977) [in Russian].
9. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 49, No. 9, 2004, pp. 1232–1235. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 127–130.

BRIEF COMMUNICATIONS

Investigation of Magnetic Properties and Micromagnetic Structure of Multicomponent $Fe_{61.4}Ni_{3.6}Cr_{3.2}Si_{2.4}Nb_{7.8}Mn_{3.6}B_{18}$ Amorphous Ribbons

E. E. Shalygina*, N. M. Abrosimova*, M. A. Komarova*, and V. V. Molokanov**

** Moscow State University, Vorob'evy Gory, Moscow, 119899 Russia*

*** Baĭkov Institute of Metallurgy and Metal Science, Russian Academy of Sciences, Moscow, Russia*

e-mail: shal@magn.ru

Received January 13, 2004

Abstract—The magnetic properties and micromagnetic structure (equilibrium distribution of magnetization) of multicomponent $Fe_{61.4}Ni_{3.6}Cr_{3.2}Si_{2.4}Nb_{7.8}Mn_{3.6}B_{18}$ amorphous ribbons are studied using scanning Kerr microscopy and a vibrating sample magnetometer. 5-mm-wide and 35-μm-thick ribbons were obtained by hardening of melt in a rapidly rotating drum. Strong difference in the surface and bulk magnetic parameters of the ribbons is established. Domain walls (DW) parallel to the ribbon length are detected. It is shown that quasistatic magnetization reversal of ribbons mainly occurs due to the DW displacement. © *2004 MAIK "Nauka/Interperiodica".*

In recent years, considerable attention has been paid to preparing and studying bulk amorphous materials [1–3]. The following bulk amorphous and nanocrystalline alloys have been recently obtained: Fe–(Al, Ga)–(P, C, B, Si) [4]; (Fe, Co, Ni)–(Zr, Hf)–B [5, 6]; and (Fe, Co, Ni)–(Zr, Nb, M)–B (M = Ti, Ta, Mo) [7, 8]. These materials were prepared at a low critical cooling rate (between 10^2 and 1 K/s) in the form of rods 1–6 mm in diameter or 1–3-mm-thick ribbons. The alloys exhibited a glass-forming ability and excellent soft properties. It was found that in obtaining bulk amorphous materials with the above sizes, it is necessary to prepare amorphous alloys at a low critical cooling rate and to have a wide temperature range of glass formation, which is determined by the difference between the crystallization (T_{X1}) and glass-formation temperature T_g ($\Delta T_X = T_{X1} - T_g$). Such properties were mainly obtained for multicomponent systems.

The object of the present study is the alloy $Fe_{61.4}Ni_{3.6}Cr_{3.2}Si_{2.4}Nb_{7.8}Mn_{3.6}B_{18}$ proposed in [9] on the basis of eutectic interaction between glass-forming phases. The glass-formation range ΔT_X determined empirically for this alloy is equal to 60 K. We will analyze here the magnetic properties and the micromagnetic structure (equilibrium distribution of magnetization) in $Fe_{61.4}Ni_{3.6}Cr_{3.2}Si_{2.4}Nb_{7.8}Mn_{3.6}B_{18}$ amorphous ribbons. This analysis allows to determine the practicability of this composition for obtaining a bulk amorphous alloy.

A 5-mm-wide and 35-μm-thick $Fe_{61.4}Ni_{3.6}Cr_{3.2}Si_{2.4}Nb_{7.8}Mn_{3.6}B_{18}$ ribbon was prepared by hardening of melt in a rapidly rotating drum. X-ray diffraction analysis revealed that the obtained ribbon is amorphous.

The near-surface magnetic properties of amorphous ribbons were studied employing a magnetooptical micromagnetometer, which allows one to measure the magnetic properties in 15–20-nm thick surface layers of the sample. A detailed description of the magnetooptical micromagnetometer can be found in [10]. The ribbons were cut into 20-mm-long pieces. An ac magnetic field of frequency f = 80 Hz was applied parallel to the sample length L. Upon the displacement of a light spot of diameter 20 μm over the ribbon surface, the distributions of the near-surface magnetization components, local magnetization curves, and hysteretic loops were measured using the equatorial Kerr effect (EKE) δ. In fact, we determined the dependences $\delta(L, H)/\delta_s \propto M(L, H)/M_s$. Here, $\delta = (I - I_0)/I_0$, where I and I_0 are the intensities of light reflected from the magnetized and nonmagnetized samples, respectively; δ_s is the EKE at $M = M_s$; and M_s is the saturation magnetization of the sample. The above measurements were made in the central part of the sample to eliminate edge effects (in particular, the variation of the local demagnetizing factor). The bulk magnetic characteristics of the sample were measured using a vibrating sample magnetometer.

Figure 1 shows (a) the bulk hysteretic loop as well as (b, c) the near-surface hysteretic loops observed on the free and wheel surfaces of the ribbon under study. Typical local magnetization curves are shown in Fig. 2. Analysis of these data leads to the following conclusions.

The ribbon studied here exhibits excellent soft magnetic properties. The bulk values of the coercive force

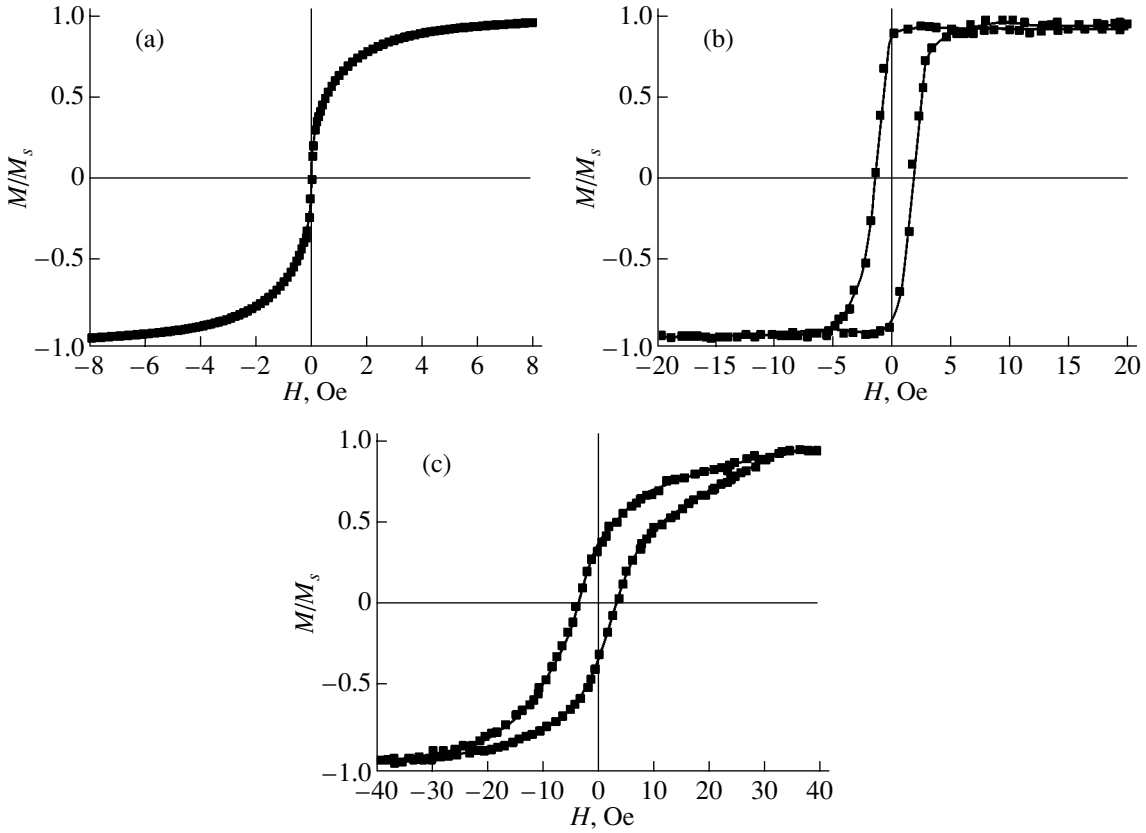


Fig. 1. Bulk hysteretic loop (a) and surface hysteretic loops observed on the free and wheel surfaces of $Fe_{61.4}Ni_{3.6}Cr_{3.2}Si_{2.4}Nb_{7.8}Mn_{3.6}B_{18}$ amorphous ribbon ((b) and (c), respectively). The magnetic field is applied along the ribbon length *L*.

H_c and saturation field H_s are equal to 0.03 and 6 Oe, respectively. The near-surface magnetic characteristics of the sample differ substantially from the bulk characteristics. The difference between the bulk and near-surface values of H_c and H_s can be explained by the presence of defects (roughness and microstructural changes) typical of surface layers of materials prepared by hardening from the melt in a rapidly rotating drum.

The values of H_c and H_s on the free surface of the ribbon are lower than on the wheel surface (1.6 and 12 Oe against 3.6 and 35 Oe, respectively). Such a difference in the near-surface magnetic properties is typical of amorphous materials obtained by hardening from the melt [10, 11]. It can be explained by the difference in the residual stresses on the wheel and free sides of the ribbon, which exists during its preparation, and by different morphologies of the surface layers.

The local near-surface magnetization curves differ insignificantly. It was found that the variations in the local values of H_c and H_s on the free and wheel surfaces of the ribbon do not exceed 6 and 10%, respectively. The high homogeneity of the local magnetic characteristics is the result of weak dispersion in the magnetic anisotropy in the sample, which in turn indicates a high quality of the multicomponent ribbon studied here.

The results of investigation of micromagnetic structures (equilibrium distribution of magnetization) deserve special attention. Figure 3 shows typical distributions of the near-surface component of the magnetization parallel to the magnetic field, which are observed on the free and wheel surfaces of the ribbon during scanning of a light spot 20 μm in diameter along the direction perpendicular to the length of the ribbon (denoted by *W*), $M/M_s(W)$. Analogous distributions of magnetization were observed for various values of *W*. The results of a more detailed analysis of the dependence $M/M_s(W)$ near the peak denoted by the arrow in Fig. 3 are depicted in Fig. 4a. Figure 4b shows the dependences of the maximal values of M/M_s of the same peak observed during the scanning of a light spot of diameter 20 μm along the ribbon length *L*. It can be seen from Fig. 3 that the dependences $M/M_s(W)$ have several periodically appearing peaks. The values of M/M_s increase with increasing *H*; the peaks are preserved, but their width increases (see Figs. 3 and 4a). It can also be seen that the values of $M/M_s^{\max}$ along the ribbon length *L* vary insignificantly (see Fig. 4b). Analogous dependences $M/M_s^{\max}(L)$ were also obtained for other peaks.

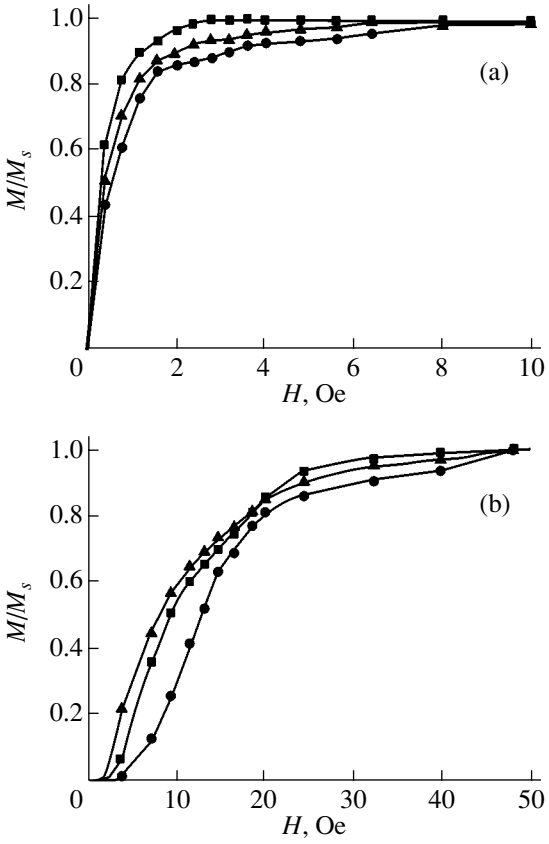


Fig. 2. Typical local near-surface magnetization curves observed on the free and wheel surfaces of the ribbon. The ac magnetic field was applied parallel to the ribbon length L ((a) and (b), respectively).

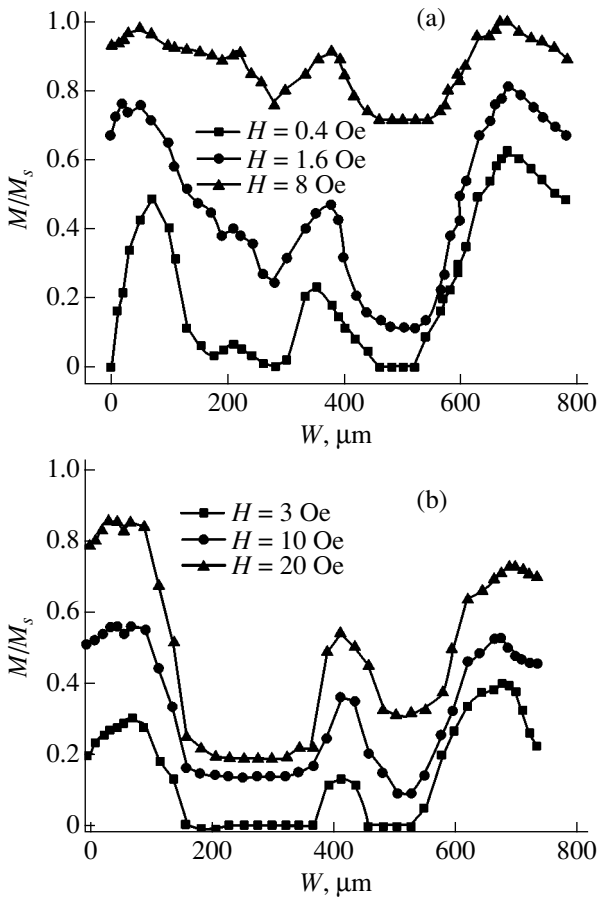


Fig. 3. Typical distributions of the magnetization components parallel to the magnetic field, observed for different values of the magnetic field ($H < H_s$) (a) on the free and (b) on the wheel sides of the ribbon. Measurements were made upon a displacement of a light spot of diameter 20 μm along a direction perpendicular the ribbon length (denoted by *W*).

Additional studies of the micromagnetic structure of the ribbon were performed using the new magnetooptical effects described in [11]. Generally speaking, the in-plane magnetization components, both parallel ($M_\parallel$) and perpendicular ($M_\perp$) to the applied magnetic field, can be measured simultaneously for any microregion being measured with the help of the EKE δ_1 and the meridional intensity effect (MIE) δ_2. The MIE is proportional to the magnetization component parallel to the plane of incidence of light ($M_\perp$ in the present case) and odd in the angle θ of deviation of the light polarization plane from the *p* component (vector **E** of the light wave is parallel to the plane of light incidence) [11]. The measurement of magnetooptical signals at $\theta = \pm 40°$ makes it possible to find the values

$$\delta^{\pm 40} = \delta_1(M_\parallel) \pm \delta_2(M_\perp). \quad (1)$$

Using this relation, we can find

$$\delta_1(M_\parallel) = (\delta^{+40} + \delta^{-40})/2, \quad (2)$$

$$\delta_2(M_\perp) = (\delta^{+40} - \delta^{-40})/2. \quad (3)$$

Analysis of magnetooptical signals carried out in [10] proved that $\delta_1(M_\perp) \neq 0$ if the magnetization reversal of the sample is executed due to rotation of the magnetization vector. It was found that the values of $\delta_1(M_\perp)$ are equal to zero for any microscopic region of the ribbon studied here.

Thus, the above results lead to the conclusion that the peaks on the dependences $M/M_s(W)$ are observed in the region of displacement of domain walls; i.e., the near-surface micromagnetic structure of the studied ribbon is characterized by the presence of domain walls parallel to the length of the ribbon. Insignificant changes in $M/M_s^{\max}$ along L are the evidence of slight distortions of domain walls in this direction. The magnetization reversal of the sample in a quasistatic magnetic field applied parallel to the ribbon length L is mainly realized due to displacement of domain walls. It should be noted that a typical feature of the magnetic-field dependence of the magnetoimpedance measured

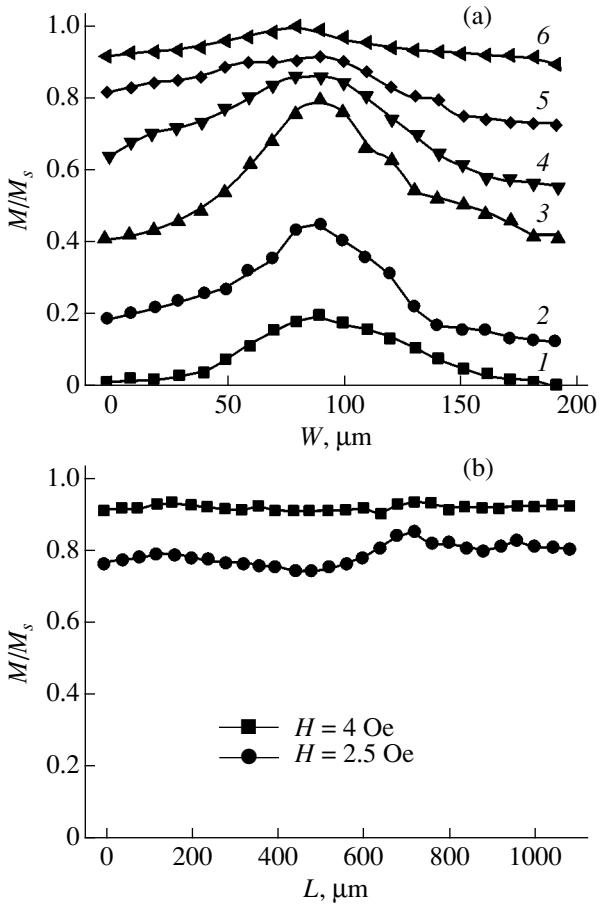


Fig. 4. (a) Distributions of the magnetization components parallel to the magnetic field observed for different values of the magnetic field ($H < H_s$) on the free side of the ribbon upon a displacement of a light spot of diameter 20 μm along the direction W perpendicular to the ribbon length near the peak marked by the arrow in Fig. 3a for $H = 0.4$ (*1*), 1.6 (*2*), 2 (*3*), 3 (*4*), 4 (*5*) and 10 Oe (*6*). (b) Magnetization distributions of the same component observed for the same peak along the ribbon length L.

on the ribbon under study in the frequency range from 100 kHz to 1 MHz is the presence of only one peak in the region of $H = 0$. In accordance with the results (see, for example, [12–15]), this fact indicates that the magnetization reversal mechanism in the ribbon, which is associated with DW displacement, is preserved up to 1 MHz.

Thus, we have analyzed the magnetostatic properties and micromagnetic structure of multicomponent $Fe_{61.4}Ni_{3.6}Cr_{3.2}Si_{2.4}Nb_{7.8}Mn_{3.6}B_{18}$ amorphous ribbons. It is found that such a ribbon is characterized by supersoft magnetic properties and a high degree of uniformity of the local surface magnetic parameters. The wide range of glass formation determined empirically for this alloy and the observed peculiarities of the magnetic properties indicate that an amorphous ribbon of this composition can be used as a precursor for preparing a bulk amorphous alloy.

ACKNOWLEDGMENTS

This work was supported financially by the Russian Foundation for Basic Research (project nos. 01-03-39986 and 02-02-16627).

REFERENCES

1. A. Inoue, Acta Mater. **48**, 279 (2000).
2. A. Inoue, A. Makino, and T. Muzushima, J. Magn. Magn. Mater. **215–216**, 246 (2000).
3. V. V. Molokanov, M. I. Petrzhik, T. N. Mikhaĭlova, and V. I. Kuznetsov, Metally, No. 5, 112 (2000).
4. A. Inoue, A. Makino, and T. Muzushima, J. Appl. Phys. **81**, 4029 (1997).
5. A. Inoue, T. Zhang, and H. Koshiba, J. Appl. Phys. **83**, 6326 (1998).
6. A. Makino, T. Bitoh, I. Murakami, *et al.*, J. Phys. (France), Part 2, 103 (1998).
7. A. Inoue, T. Zhang, H. Koshiba, and T. Itoe, Mater. Res. Soc. Symp. Proc. **554**, 251 (1999).
8. H. Chiriac and N. Lupu, J. Magn. Magn. Mater. **215–216**, 394 (2000).
9. V. V. Molokanov, A. N. Shalygin, M. I. Petrzhik, *et al.*, Perspekt. Mater., No. 3, 10 (2003).
10. E. Shalyguina, L. Bekoeva, and K. H. Shin, J. Magn. Magn. Mater. **215–216**, 472 (2000).
11. E. Shalyguina, L. Bekoeva, and N. Tsidaeva, Sens. Actuators **81**, 216 (2000).
12. L. V. Panina and K. Mohri, J. Magn. Magn. Mater. **157–158**, 137 (1996).
13. L. V. Panina, K. Mohri, T. Uchiyama, and M. Noda, IEEE Trans. Magn. **31**, 1249 (1995).
14. F. L. A. Machado, C. S. Martins, and S. M. Rezende, Phys. Rev. B **51**, 3926 (1995).
15. K. Mohri, K. Bushida, M. Noda, *et al.*, IEEE Trans. Magn. **31**, 2455 (1995).

Translated by N. Wadhwa

Technical Physics, Vol. 49, No. 9, 2004, pp. 1236–1239. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 131–133.

BRIEF COMMUNICATIONS

Removal of Pyramidal Germanium Nanoislands from the Silicon (001) Surface as a Result of Deposition/Sputtering of a Gold Nanolayer

A. I. Stogniĭ*, N. N. Novitskiĭ, O. M. Stukalov**, A. I. Demchenko*, and V. I. Khit'ko***

** Minsk Research Institute of Radio-Engineering Materials, Minsk, Belarus*

e-mail: stognij@ifttp.bas-net.by

*** Institute of Solid-State and Semiconductor Physics, Belarussian Academy of Sciences, ul. Brovki 17, Minsk, 220072 Belarus*

e-mail: novitski@ifttp.bas-net.by

Received January 13, 2004

Abstract—Optimal conditions for selective removal of pyramidal germanium nanoislands from the silicon (001) surface by sequential ion-beam deposition/sputtering of a gold layer with a nanoscale thickness are determined experimentally; it is important that the sizes of large domelike germanium nanoislands change only slightly as a result of the ion-beam treatment. *© 2004 MAIK "Nauka/Interperiodica".*

Device-oriented semiconductor heterostructures with germanium nanoislands must have stable properties, perfect crystal structure, and uniform distribution of nanoisland sizes [1]. However, the germanium nanoislands on the silicon (001) surface have the shape of pyramids with differing sizes or of larger hemispherical domes [2]. In addition, the smallest germanium islands can be elongated and have irregular shapes (huts) [3]. The diversity of types of nanoislands is caused by the effect of a number of controlled and uncontrolled factors on the conditions of growth of heterostructures and on relaxation of stresses. In this context, it appears important to use postgrowth methods in order to obtain a more uniform distribution of nanoisland sizes. For example, using the fact that the smallest domes exceed in size the largest pyramids [2], one could employ ion-beam sputtering in order to remove the smallest islands. However, further studies of the effect of irradiation with ions on the properties of islands are needed here. It should also be taken into account that, as a result of sputtering, the germanium nanoislands can become contaminated with impurity particles, in particular, those of silicon. Therefore, the use of the method based on the ion-beam deposition/sputtering of a nanolayer of material whose sputtering rate is comparable to that of hillocks in the profile of the starting surface appears to be more promising. This method was used previously for subnanometer-scale polishing of optical surfaces [4]. In the case under consideration, we are dealing with surface hillocks in the form of germanium nanoislands; as a result, gold is a suitable material for a smoothening nanolayer since gold is chemically inert and has the rate of sputtering by argon ions that almost coincides with the corresponding rate for germanium [5]. The method under consideration includes (i) the preferential filling of valleys in the surface profile at the first stage of ion-beam deposition of a gold nanolayer on the starting surface and (ii) the ion sputtering of this layer (at the second stage) together with germanium nanoislands whose sizes are smaller than the nanolayer thickness. Since the rate of argon-ion sputtering of silicon is lower by at least a factor of 2 than those of germanium and gold [5] and the gold layer shades the silicon surface from ion bombardment, contamination of the surface by silicon can be minimized experimentally.

The setup for combined ion-beam deposition/sputtering was similar to that described previously [4]. The operating gas was argon. The gold flux was normal to the surface; the deposition rate was 7 nm/min. The argon ions for sputtering had an energy lower than 350 eV; the ion-beam current density was equal to 0.06 mA/cm^2. The sputtering rate for gold was lower than 1.2 nm/min. The rates of deposition and sputtering of gold were determined using flat areas of quartz-substrate surface (see [6]). The germanium nanoislands on the silicon (001) surface were formed by molecular-beam epitaxy. The method for forming an array of nanoislands was described in detail by Leifeld *et al*. [3]. The silicon wafer was kindly placed at our disposal by D. Grützmacher. The samples were neighboring parts cleaved from the wafer. We analyzed the surface morphology using a Femtoskan-001 atomic-force micro-

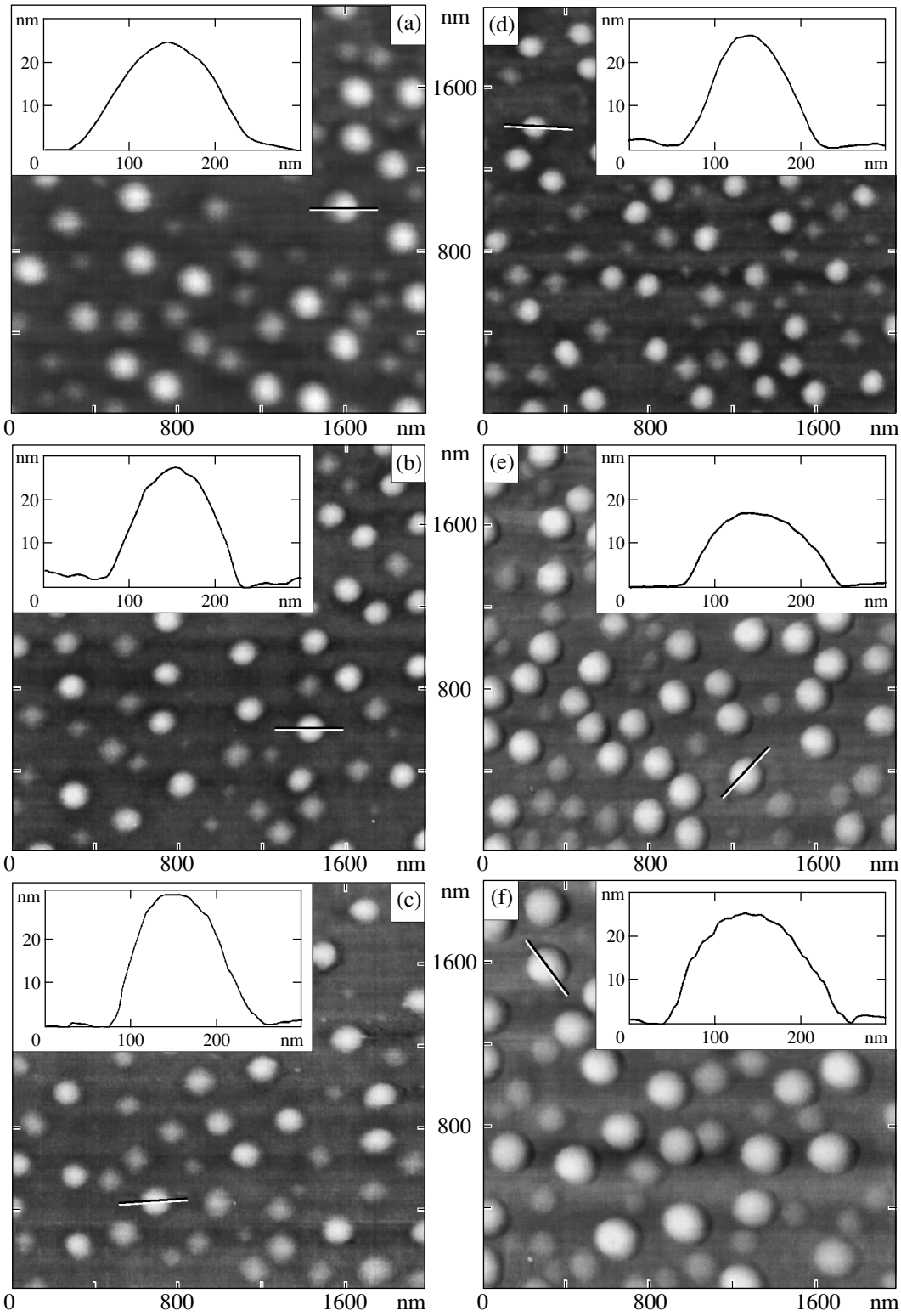


Fig. 1. The AFM images of nanosized germanium islands (a) on the initial silicon surface, (b–d) after deposition of the gold nanolayer, and (e, f) after sputtering with argon ions.

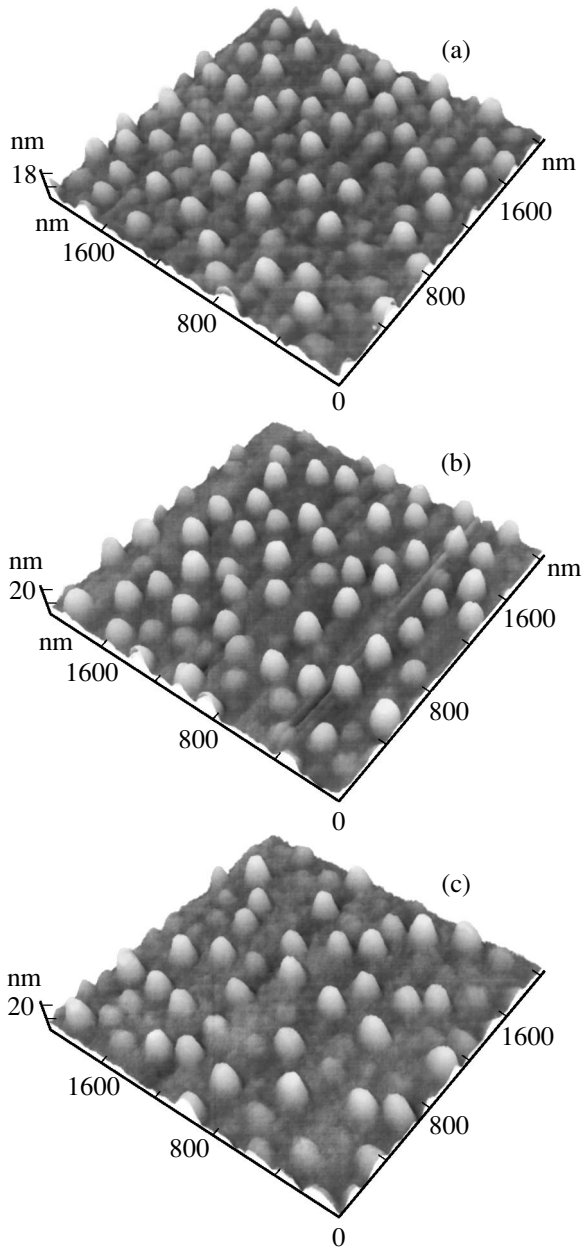


Fig. 2. Selective sputtering of pyramidal germanium nanoislands.

scope (AFM) produced by the Center for Promising Technologies (Moscow State University). The microscope operated in the contact mode; CSC12 silicon cantilevers produced by MicroMasch Co. (www.spmtis.com) were used; and the scanned area was as large as $5 \times 5\ \mu m^2$.

In Fig. 1, we show the AFM images of the (a) initial surface; the surfaces after deposition of the gold layer for (b) 5, (c) 10, and (d) 20 s; (e) the surface after deposition of the gold layer for 5 s with subsequent sputtering for 4 min; and (f) the surface after deposition of the gold layer for 10 s with subsequent sputtering for 4 min. Cross sections of characteristic nanosized germanium islands seen in Fig. 1 are shown in the insets. Evidently, the AFM data are insufficient for unambiguous consideration of how the surface profile evolves after each stage of the deposition/sputtering of the gold nanolayer. However, the set of AFM images as a whole indicates that gold is deposited in the form of sharp tips on domes and thin layers over the lateral faces of pyramids; these layers are thicker at the top and thinner at the pedestal. Simultaneously, the pyramids are buried as the thickness of the gold layer increases at the surface areas free from germanium islands (Figs. 1b–1d). Irradiation with ions for 4 min is found to be sufficient for sputtering off of the gold layer deposited for 5 s and for the onset of sputtering of germanium nanoislands (Fig. 1e). At the same time, the gold layer deposited for 10 s is removed only partially as a result of irradiation with ions for 4 min (Fig. 1f). This inference is confirmed by the results of measuring the electrical conductance and optical absorption in the case of quartz substrates. It can be also seen from Fig. 1 that the dome-shaped germanium islands broaden as a result of ion sputtering (see insets to Figs. 1e, 1f). In this case, the processes of deposition and sputtering are characterized by the duration of the process rather than by the corresponding rates as a result of variations in these rates at the surface with a complex profile for the time intervals under consideration. According to Fig. 1, removal of the smallest germanium islands is observed after deposition of the gold layer for more than 10 s and subsequent sputtering with argon ions for more than 4 min. In Fig. 2, we show the AFM images of the surfaces that were sputtered for 8 min after deposition of the gold layer for (a) 10, (b) 20, and (c) 30 s. Figure 2a indicates that the pyramids are completely sputtered off, the domes are partially sputtered, and a complex profile is formed in the surface areas free of germanium islands as a result of partial sputtering of the surface after removal of the masking gold layer. It can be seen from Fig. 2b that only the traces of pyramids are observed after sputtering, whereas the dome-shaped germanium nanoislands remain close in shape and size to the domes on the initial surface (Fig. 1a). Pyramids and domes seen in Fig. 2c are found to be partially buried under incompletely removed gold layer. This behavior was to be expected since the thickest attainable gold layer (about 4 nm thick) was deposited for 30 s.

Thus, the experimentally optimized conditions of ion-beam deposition/sputtering of the gold nanolayer make it possible to remove selectively the smallest germanium nanoislands from the silicon surface, including those with pyramidal shape; larger, dome-shaped nanoislands remain intact. This circumstance leads to a substantial increase in the uniformity of size distribution of germanium nanoislands. The increased uniformity is of much interest with respect to the problem of

formation of smooth surfaces with nanohillocks of certain sizes.

ACKNOWLEDGMENTS

In conclusion, we thank D. Grützmacher (Laboratory for Micro- and Nanotechnology, Paul Scherer Institute, Switzerland) for placing at our disposal the samples of operating structures and for constructive criticism.

REFERENCES

1. N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 385 (1998) [Semiconductors **32**, 343 (1998)].
2. G. Medeiros-Ribeiro, A. M. Bratkoski, Th. I. Kamins, *et al.*, Science **279**, 353 (1998).
3. O. Leifeld, R. Hartmann, E. Müller, *et al.*, Nanotechnology, No. 10, 122 (1999).
4. A. I. Stognič, N. N. Novitskič, and O. M. Stukalov, Pis'ma Zh. Tekh. Fiz. **28** (1), 39 (2002) [Tech. Phys. Lett. **28**, 17 (2002)].
5. V. S. Smentkowski, Prog. Surf. Sci. **64**, 1 (2000).
6. A. I. Stognič, N. N. Novitskič, S. D. Tushina, and S. V. Kalinnikov, Zh. Tekh. Fiz. **73** (6), 86 (2003) [Tech. Phys. **48**, 745 (2003)].

Translated by A. Spitsyn

Technical Physics, Vol. 49, No. 9, 2004, pp. 1240–1242. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 134–136.

BRIEF COMMUNICATIONS

Special Correcting Winding for a Torsatron

V. G. Kotenko

Kharkov Institute of Physics and Technology National Science Center, Akademicheskaya ul. 1, Kharkov, 61108 Ukraine

e-mail: volkov@ipp.kharkov.ua

Received February 4, 2004

Abstract—The possibility of controlling in two dimensions the position of the region of closed magnetic surfaces in a helical magnetic confinement system is demonstrated by numerical simulations carried out using a two-pole model torsatron as an example. The displacement of this region in a direction perpendicular to the equatorial plane of the torus is provided by the magnetic field produced by a special correcting winding. *© 2004 MAIK "Nauka/Interperiodica".*

The first experimental investigations of the magnetic field structure in a closed helical magnetic confinement system (the Liven'-1 stellarator) by high-resolution methods [1] found definite indications that the region of closed magnetic surfaces is always displaced in a direction perpendicular to the equatorial plane of the torus (in the axial direction along the symmetry axis of the torus). The displacement was explained as being due to the magnetic field perturbations, which may result from the imperfect construction and assemblage of the individual elements of a magnetic coil winding. In [1], experimental measurements were also carried out on the amount by which the region of closed magnetic surfaces is displaced in the radial direction (i.e., along the major radius of the torus in its equatorial plane) under the action of a comparatively weak uniform transverse magnetic field. These measurements demonstrated that the radial position of the plasma column in stellarator devices can be controlled by means of correcting ring coils. However, the question about the possibility of correcting the axial component of the displacement of the region of closed magnetic surfaces was not addressed in that paper and still remains open. That is why the means for controlling the position of this region in two dimensions are absent in the existing helical and helical-like toroidal magnetic confinement systems and are not considered in the design of the systems currently under development.

Numerical simulations described in the present paper demonstrate that the axial displacement of the region of closed magnetic surfaces in a torsatron can be controlled by a special correcting winding. The geometry of the winding indicates what kind of imperfection in the construction of a magnetic coil winding is responsible for the axial displacement of the region of closed magnetic surfaces.

The special correcting winding consists of auxiliary helical coils wound in a certain manner around the torus carrying the main helical winding of the torsatron. Each of the l poles of the main helical winding is associated with one pole of the auxiliary helical winding. Each point of the auxiliary base helical line along which the conductor of a pole of the auxiliary winding is wound on the torus is at a distance S = const from the main base line for the corresponding pole of the main helical winding. The distance S = const is measured from the parallel passing through this point on the torus. In Cartesian coordinates such that the z axis is aligned with the symmetry axis of the torus, the equation of the auxiliary line can be written in the following parametric form, which is convenient for numerical calculations:

$$x = (R_0 + a\cos(\theta(\varphi)))\cos(\varphi \pm S/(R_0 + a\cos(\theta(\varphi)))),$$

$$y = (R_0 + a\cos(\theta(\varphi)))\sin(\varphi \pm S/(R_0 + a\cos(\theta(\varphi)))), \quad (1)$$

$$z = a\sin(\theta(\varphi)).$$

Here, R_0 and a are the major and minor radii of the torus, θ and φ are the poloidal and toroidal angles, $\theta(\varphi)$ is the winding law for the main helical winding, and the sign in the argument of the trigonometric functions is chosen according to the direction along which the distance S is measured. Numerical simulations were carried out for a model torsatron shown schematically in Fig. 1. The torsatron magnetic system includes a two-pole (l = 2) main helical winding wound along base helical lines *1* and *2* on a torus whose radii are in the ratio a/R_0 = 0.25. The magnetic system should also involve at least one pair of compensating ring-shaped coils *3* with built-in correcting ring-shaped coils (which are not shown in the figure). The compensating coils are

1063-7842/04/4909-1240$26.00 © 2004 MAIK "Nauka/Interperiodica"

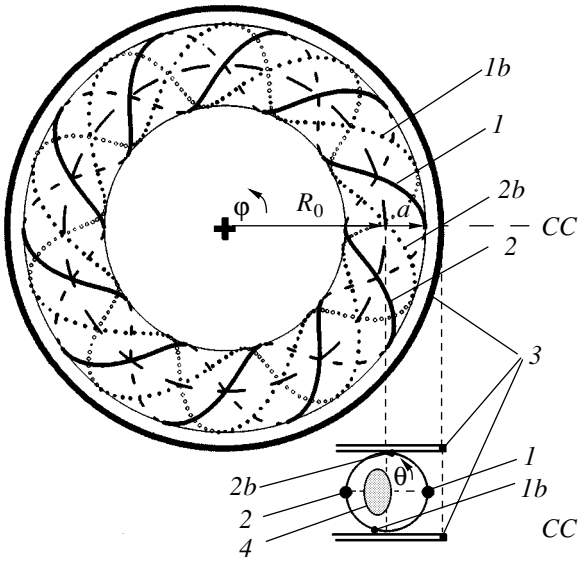


Fig. 1. Schematic of a model torsatron for numerical simulations: top view and view in cross section *CC*.

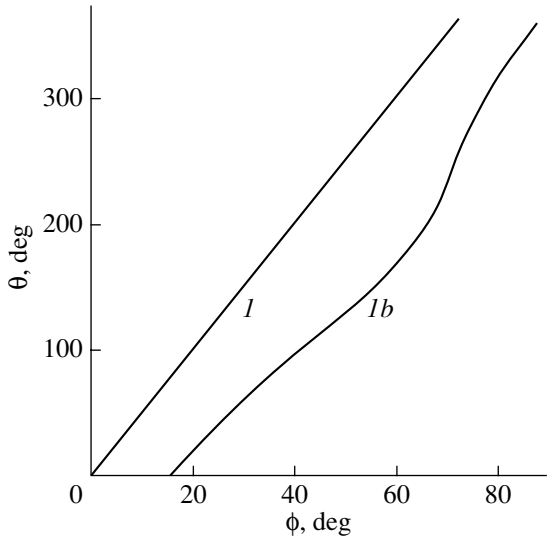


Fig. 2. Dependence of the poloidal angle θ on the toroidal angle φ along main base helical line *1* and auxiliary base helical line *1b* over one pitch length.

required to produce the region of closed magnetic surfaces *4* within the volume enclosed by the helical winding. The correcting ring-shaped coils serve to control the radial position of this region in the equatorial plane of the torus.

The winding law of base helical lines *1* and *2* is given by a cylindrical helix: $\theta(\varphi) = m\varphi$, where $m = 5$ is the number of pitch lengths of each of the lines over the full length of the torus along the toroidal axis. Base helical lines *1b* and *2b* of the auxiliary helical winding correspond, respectively, to base lines *1* and *2* of the main helical winding. The points of intersection of lines *1* and *1b* (*2* and *2b*) with the same parallel on the torus are separated by a distance of $S/R_0 = 0.35$. This distance is measured from line *1* (*2*) along the parallel in the direction in which the toroidal angle φ is increasing.

For the model torsatron under consideration, Fig. 2 graphically illustrates the winding laws of main base line *1* and auxiliary base helical line *1b*. The difference of the "derived" winding law of the auxiliary helical winding from the "original" linear winding law of the main helical winding is seen to be highly nonlinear and is very difficult to describe analytically.

Figure 3 presents the calculated cross sections of the closed magnetic surfaces in the model torsatron in question. The cross sections shown in Fig. 3a refer to the case in which the current in the auxiliary winding is switched off. These cross sections are characteristic of a conventional torsatron with a planar magnetic axis: the region of closed magnetic surfaces is seen to be symmetric about the equatorial plane of the torus. Figure 3b refers to the case in which the current in the poles of the auxiliary winding is equal to 0.03 of the current in the poles of the main helical winding. In this case, all the cross sections of the magnetic surfaces are seen to be displaced upward from the equatorial plane by an amount of about $0.1a$. For an oppositely directed current in the auxiliary winding, the magnetic surfaces are displaced downward. The same effect is captured by calculations for the case in which the current in the auxiliary winding is not reversed but the distance *S* is measured in the opposite direction (or in the direction opposite to that of the angular coordinate φ). Simulations also show that the magnetic field of the auxiliary helical winding does not change the radial position of the region of closed magnetic surfaces in the equatorial plane of the torus.

Hence, the auxiliary helical winding that has been considered above is fully capable of functioning as a special correcting winding for controlling the axial position of the region of closed magnetic surfaces. In combination with an ordinary ring-shaped correcting winding, it can be used, if necessary, to control in two dimensions the position of the plasma column in closed magnetic confinement systems of the stellarator type.

In conclusion, it is important to point out the following circumstance. If the conductor of a pole of, e.g., a single-layer main helical winding is wrapped beginning with the base line on one side of it in a turn-by-turn fashion, the thickness of the insulator between the adjacent turns being constant, then the helical winding law of the last turn in the layer may be similar to Eq. (1) with the parameter $S = (n - 1)d$, where *n* is the number of turns in the layer and *d* is the diameter of the insu-

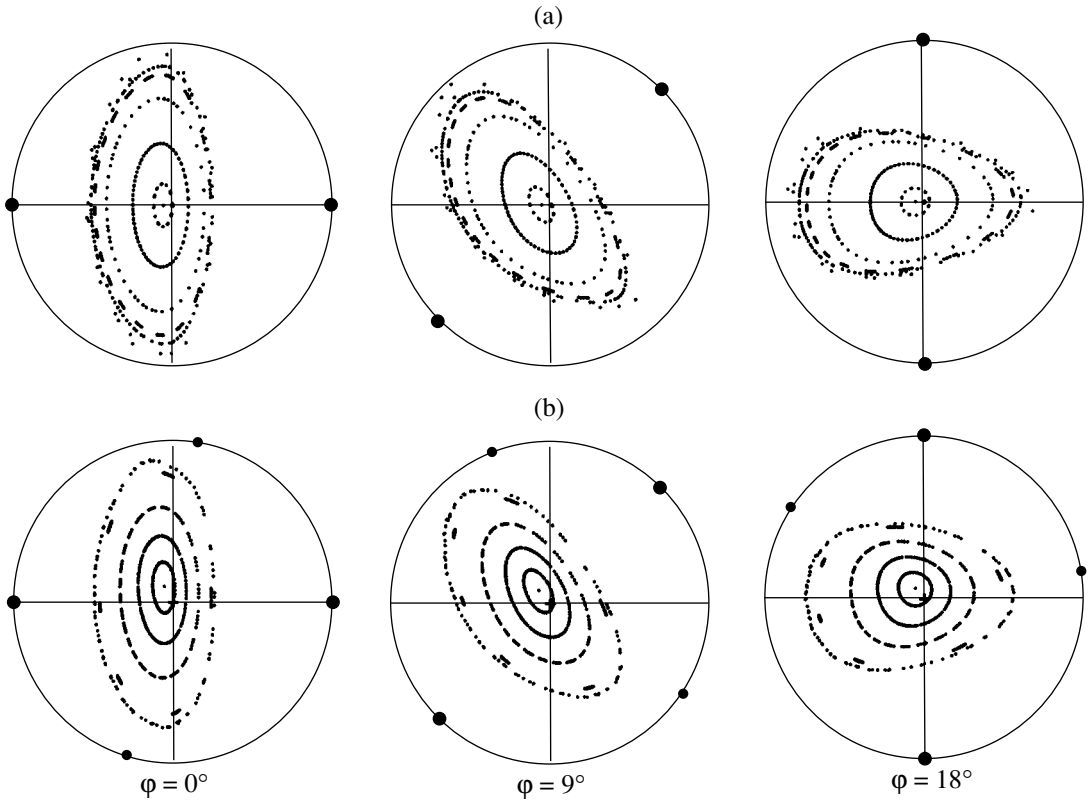


Fig. 3. Cross sections of the magnetic surfaces over a half-period of the magnetic field when (a) the current in the special correcting winding is switched off and when (b) the current in the special correcting winding is equal to 0.03 of the current in the main helical winding.

lated conductor. As a result, the calculated position of the region of closed magnetic surfaces in a magnetic confinement system can differ from its actual position. This problem can be partially resolved by using a symmetrized pole whose conductor is wrapped beginning with the base line on both sides of it or the turns of whose conductors are wound according to the winding law of the base helical line.

REFERENCES

1. M. S. Berezhetskiĭ, S. E. Grebenshchikov, A. P. Popryadukhin, and I. S. Shpigel', Zh. Tekh. Fiz. **35**, 2167 (1965) [Sov. Phys. Tech. Phys. **10**, 1662 (1965)].

Translated by O. Khadin

Technical Physics, Vol. 49, No. 9, 2004, pp. 1243–1246. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 137–140.

BRIEF COMMUNICATIONS

Growth and Photoelectric Properties of Graded-Gap Si–$(Si_2)_{1-x}(GaP)_x$ Heterostructures

B. Sapaev, A. S. Saidov, and S. Dadamukhamedov

Physicotechnical Institute and the Research and Production Association "Physics–Sun," Academy of Sciences of Uzbekistan, Tashkent, 700084 Uzbekistan

e-mail: amin@physic.uzsci.net

Received February 9, 2004

Abstract—Theoretical grounds for formation of continuous substitutional solid solutions are analyzed taking into account the generalized moments, the difference in valence, and covalent radii of initial components. On the basis of these studies, the technology of formation of epitaxial $(Si_2)_{1-x}(GaP)_x$ $(0 \le x \le 1)$ layers on silicon substrates from the tin solution–melt using forced cooling is developed. The distribution of components over the thickness of Si–$(Si_2)_{1-x}(GaP)_x$ layers, the photosensitivity, and the current–voltage characteristics of the Si–$(Si_2)_{1-x}(GaP)_x$ heterostructures are studied. Analyses of the results of the X-ray studies and photoelectric properties of obtained solid-solution epitaxial layers indicate that the grown graded-gap $(Si_2)_{1-x}(GaP)_x$ layers have a high structural quality. *© 2004 MAIK "Nauka/Interperiodica".*

Research in the field of present-day materials science of semiconductors is aimed at production of new semiconductor materials that have high quality and exhibit the improved characteristics compared to those of the materials used currently in photoelectronics. Therefore, the development of technology and the study of properties of obtained new semiconductor materials and solid solutions based on these materials are of utmost importance in this field. In this context, heteroepitaxial deposition of GaP onto Si substrates is of much interest since the lattice parameters of Si and GaP are almost the same, whereas the band gaps are radically different. Substitutional solid solutions of Si and GaP can be formed since the conditions $\Delta Z = 0$ and $|\Delta r| = (r_A + r_B) - (r_C - r_D) \le 0.1$ [1] are satisfied for these solutions (here, Z is the valence and r are the covalent radii of the components).

It was shown for the first time by Alferov *et al.* [2] by the example of a $(Ge_2)_{1-x}(GeAs)_x$ system that pyrolytic synthesis from gaseous phase can be used to form a continuous sequence of solid solutions. We also previously showed [3] that it is possible to obtain graded-gap $(Ge_2)_{1-x}(GeAs)_x$ $(0 \le x \le 1)$ metastable solid solutions on the Ge and GaAs using a Pb solution–melt as the source. An analysis of solubility of Si and GaP in liquid metallic solvents show that the use of tin as the solvent and a relatively low growth temperature (750–900°C) are conducive to the formation of $(Si_2)_{1-x}(GaP)_x$ solid solutions.

In this paper, we report the results of studies related to the technology and morphology of Si–$(Si_2)_{1-x}(GaP)_x$ heterostructures obtained from a limited volume of a Sn solution–melt in the atmosphere of hydrogen purified with palladium; we also report the results of studying the certain photoelectric properties of these heterostructures. A total of 15 Si–$(Si_2)_{1-x}(GaP)_x$ heterostructures with an area $S \approx 15$ mm^2 each were fabricated and studied. Technological processes related to fabrication of heterostructures from a limited volume of the solution–melt were described by Saidov *et al*. [4].

Epitaxial layers with the thickness in the range from 15 to 30 μm had the *n*-type conductivity. According to the data on the distribution of components over the thickness obtained using a Cameca electron-probe analyzer, the GaP content in the graded-gap epitaxial $(Si_2)_{1-x}(GaP)_x$ layer increases along the growth axis and is as high as 48% for Ga, 52% for P, and 0% for Si (Fig. 1a). The raster patterns obtained using a Jeol JSM 5910 LV X-ray microanalyzer (Fig. 1b) show that there are no macroscopic structural defects and metallic inclusions of the second phase. The measurement error was no larger than 2%.

We used a DRON-3M diffractometer to determine the interplanar spacings d_{nkl} of atoms on the substrate side and on the side of epitaxial layers. Using the formula $a = d\sqrt{h^2 + k^2 + l^2}$, where h, k, and l are the Miller indices, we determined the lattice constants for Si and $(Si_2)_{1-x}(GaP)_x$ solid solutions: $a_{Si} = 5.4290$ Å and $a_{(Si_2)_{1-x}(GaP)_x} = 5.4293$ Å for $x \approx 0.5$. The calculated error in determining the lattice parameters a is equal to $\Delta a = 0.0004$ Å. The smooth and graded-gap $(Si_2)_{1-x}(GaP)_x$ epitaxial layers are obtained using forced cooling with a rate of 0.5–5 K/min at the distance $\delta = 0.75$–1.0 mm between two horizontally arranged Si substrates.

1063-7842/04/4909-1243$26.00

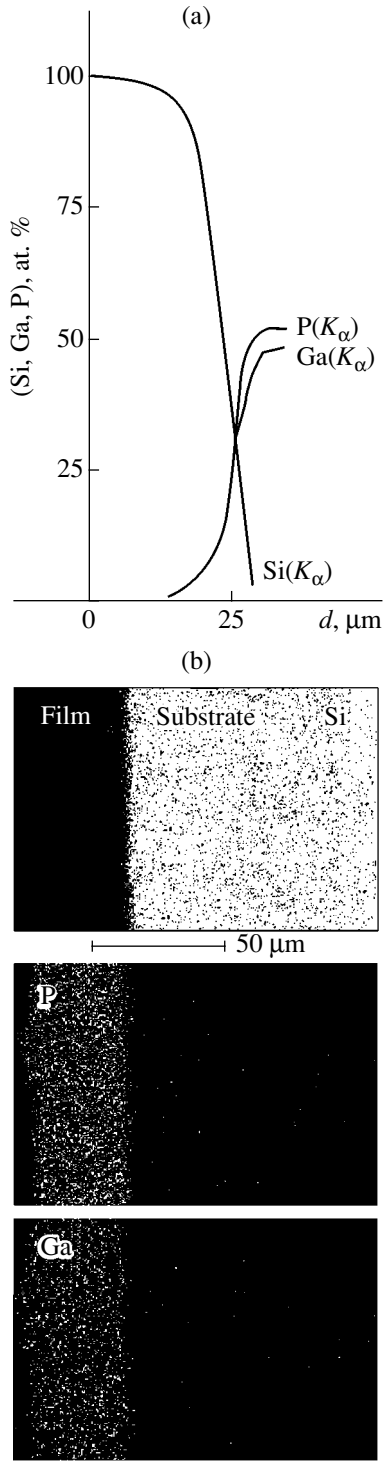


Fig. 1. (a) The distribution of the Ga_{K_α}, P_{K_α}, and Si_{K_α} components over the thickness of Si–$(Si_2)_{1-x}(GaP)_x$ structures and (b) the raster patterns of the cleaved surface of the Si–$(Si_2)_{1-x}(GaP)_x$ structure.

As the GaP content in the epitaxial $(Si_2)_{1-x}(GaP)_x$ layers increases and for certain thicknesses of these layers, the role of the thermal-expansion coefficient increases owing to a gradual transition from Si to GaP. As a result, we do not observe any bends or fractures in the $(Si_2)_{1-x}(GaP)_x$ films.

In the course of forced cooling, the heat removal from the last crystallizing layers proceeds more rapidly than that from initial layers as a result of nonuniform heat removal from different sites. In addition, as the cooling rate increases, this difference in the heat removal increases; as a result, thermoelastic stresses appear. In turn, these stresses give rise to the plastic strain, formation of cracks, and even destruction of the film. The causes of formation of similar defects were outlined previously [5]. In authors' opinion, the causes of formation of the defects are related to the following factors: a difference between the crystal-lattice parameters, thermal stresses, the composition gradient over the epitaxial-layer thickness, and introduction of defects from the substrate.

Mismatch of lattice parameters for Si–GaP heterostructures amounts to 0.36% and, thus, is insignificant. Therefore, the effect of stresses arising at the interface between substrate and epitaxial layer owing to the difference between the lattice constants in the heterostructure does not exist; it is noteworthy that the fraction of the GaP chemical component varies gradually from zero to unity ($0 \le x \le 1$). Introduction of defects from the substrate is eliminated by choosing dislocation-free Si substrates of high structural quality. The cracking of Si layers grown on Ge was observed in a Ge–Si structure [6, 7]; cracking of thick ZnSe layers grown on GaAs was also observed [8]. Stresses arising in the film owing to the difference in the thermal-expansion coefficients can be estimated using the formula [9]

$$\sigma_{\Delta\alpha} = \frac{E}{1-\gamma}\Delta\alpha\Delta T,$$

where E is Young's modulus, γ is the Poisson ratio, $\Delta\alpha$ is the difference between the thermal-expansion coefficients of epitaxial film and the substrate, and ΔT is the difference between the temperature at which the film is grown and room temperature.

As temperature decreases, the stress $\sigma_{\Delta\alpha}$ increases linearly, whereas the plasticity increases exponentially. As a result, the relaxation of thermal stress is hampered, which leads to the fracture of epitaxial layers [9].

We performed preliminary studies of electrical properties of fabricated Si–$(Si_2)_{1-x}(GaP)_x$ structures. Ohmic contacts were formed using a Ga–In alloy. The dark current–voltage (I–V) characteristics of p-Si–n-$(Si_2)_{1-x}(GaP)_x$ heterojunctions were measured at both forward and reverse bias voltages V_b at the temperature T = 290 K (Fig. 2a). There are numerous theories that describe the I–V characteristics for homojunctions and heterojunctions [10]. It is well known that electrical characteristics of a forward-biased p–n junction are independent of both the potential-barrier height V_d and the resistivity of semiconductors. Irrespective of the theory chosen for description of I–V characteristics, the

forward current J_{dir} increases exponentially with increasing voltage V_b and follows the law $J_{dir} \propto \exp(qV_b/kT) - 1$. The potential barrier V_d determined from extrapolation of the linear portion of the characteristic is equal to 0.5 V. In the case of $V_d = V_b = 0.5$ V, the barrier disappears and further increase in current with increasing V_b is governed by physical processes that occur in the semiconductor bulk (thermal emission and tunneling).

It is easy to recognize three specific portions in the reverse *I–V* characteristic: (i) at the voltage range V_b = 0–0.3 V, (ii) at V_b = 0.3–1.25 V, and (iii) at $V_b > 1.25$ V. In the region of low voltages ($V_b < 0.3$ V), the reverse current increases only slightly with increasing voltage, follows approximately the law $J = B \times V_b^n$, and is governed predominantly by tunneling current [10] (B is a constant and $n < 1$). In the second voltage range (V_b = 0.3–1.25 V), a significant increase in the reverse current is observed. At still higher voltages (V_b = 2.1 V), the impact ionization sets in the bulk of the *p–n* heterojunction, which eventually leads to the soft breakdown. Apparently, the increase in J in the voltage range V_b = 0.3–1.25 V is related to the development of impact-ionization processes in the regions adjoining the *p–n* heterojunction rather than in the bulk. This circumstance is caused by variation in the depletion-region width that is governed by the charge of impurity atoms in the aforementioned regions. Thus, the reverse current at V_b = 0.3–1.25 V is mainly governed by the onset of impact ionization of impurity centers in the regions that adjoin the *p–n* heterojunction.[1]

In Fig. 2b, we show a typical spectral dependence of photocurrent in pSi–$n(Si_2)_{1-x}(GaP)_x$ structures. The heterojunctions were irradiated with light perpendicularly to the plane of *p–n* junction; the graded-gap epitaxial layer was exposed to light. As can be seen, the photosensitivity of the structures is observed in the photon-energy range E_{ph} = 1.05–2.4 eV.

The short-wavelength edge of spectral sensitivity is governed by the epitaxial layer of the graded-gap $(Si_2)_{1-x}(GaP)_x$ crystal. The crystal composition was chosen such that the band gap increased from the side of the Si substrate and was as wide as possible at the crystal surface that served as the wide-gap entrance window.

The *n–p* junction is deep in the structures under consideration (recall that the thickness of the *n*-type graded-gap layers $d \approx$ 15–30 μm); therefore, the efficiency of collection of photogenerated charge carriers is controlled by the graded-gap field E_V in the $(Si_2)_{1-x}(GaP)_x$ crystal. The charge carriers generated by the short-wavelength radiation near the surface of the wide-gap entrance window are collected by the field E_V with an efficiency higher than that for the charge carriers photogenerated in the crystal bulk. This inference is confirmed by the presence of a broad band in the spectral dependence of photocurrent in the photon-energy range E_{ph} = 1.35–2.1 eV. A gradual increase in the photoresponse intensity and its peak at E_{ph} = 2.05 eV are caused by an increase in the value of E_V with approach to the surface of the $(Si_2)_{1-x}(GaP)_x$ crystal due the gradient in the band gap. The sharp short-wavelength falloff of the spectral characteristic ($\lambda <$ 0.5–0.6 μm) is a result of the surface recombination of photogenerated charge carriers, whereas the long-wavelength falloff (at λ = 0.918–1.180 μm) is caused by recombination of charge carriers in the bulk of the graded-gap $(Si_2)_{1-x}(GaP)_x$ crystal.

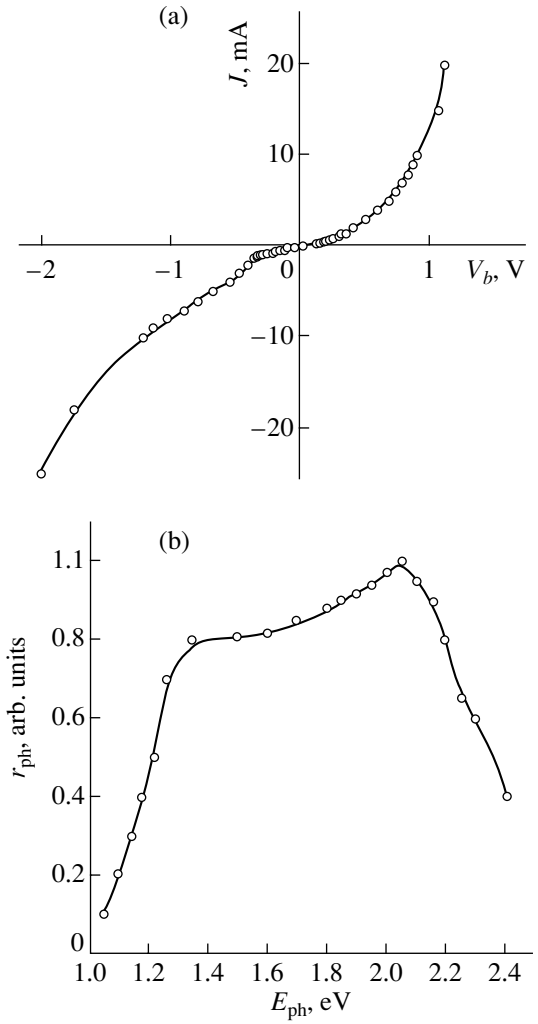


Fig. 2. (a) The dark *I–V* characteristic and (b) the spectral dependence of photoresponse of the pSi–$n(Si_2)_{1-x}(GaP)_x$ heterostructure.

To summarize, we fabricated new graded-gap $(Si_2)_{1-x}(GaP)_x$ solid solutions on Si substrates. These

[1] In this paper, we do not intend to study in detail the physical mechanisms responsible for the behavior of the *I–V* characteristics. The results of corresponding studies will be reported in our subsequent publication.

solid solutions exhibit a wider range of spectral sensitivity compared to the graded-gap $Al_xGa_{1-x}As$ solid solutions that are widely used for fabrication of solar cells and are grown exclusively on expensive GaAs substrates.

Thus, the above consideration suggests that new solid solutions on silicon substrates could be used widely in photoelectronics.

REFERENCES

1. M. S. Saidov, Geliotekhnika, Nos. 5–6, 57 (1997).
2. Zh. I. Alferov, M. Z. Zhingarev, S. G. Konnikov, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **16**, 831 (1982) [Sov. Phys. Semicond. **16**, 532 (1982)].
3. A. S. Saidov, É. A. Koshchanov, B. Sapaev, *et al.*, Dokl. Akad. Nauk Resp. Uzb., No. 2, 26 (1988).
4. A. S. Saidov, A. Kutlimratov, B. Sapaev, *et al.*, Vestn. Gulistansk. Gos. Univ., No. 2, 40 (2001).
5. M. G. Mil'vidskiĭ and V. V. Osvenskiĭ, *Structural Defects in Epitaxial Semiconductor Layers* (Metallurgiya, Moscow, 1985) [in Russian].
6. A. K. Riben, B. L. Feucht, and W. G. Oldham, J. Electrochem. Soc. **113**, 245 (1966).
7. H. J. Howel and A. G. Milnes, J. Electrochem. Soc. **116**, 843 (1969).
8. A. G. Milnes and D. L. Feucht, *Heterojunctions and Metal–Semiconductor Junctions* (Academic, New York, 1972; Mir, Moscow, 1975).
9. V. M. Ustinov and B. G. Zakharov, Obz. Élektron. Tekh., Ser. 6: Materialy, No. 4 (1977).
10. B. L. Sharma and R. K. Purohit, *Semiconductor Heterojunctions* (Pergamon, Oxford, 1974; Sov. Radio, Moscow, 1979).

Translated by A. Spitsyn

Technical Physics, Vol. 49, No. 9, 2004, pp. 1247–1248. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 9, 2004, pp. 141–142.

BRIEF COMMUNICATIONS

On the Interpretation of Optical Accelerometer Indications

A. S. Devyatisil'nyĭ

Institute of Automatics and Control, Far-East Division, Russian Academy of Sciences, Vladivostok, 690041 Russia

e-mail: devyatis@iacp.dvo.ru

Received February 10, 2004

Abstract—A basic scheme of "meters" involved in an optical accelerometer based on the effect of relativistic frequency shift is presented. The point-spread function of the instrument is estimated in terms of both conventional and advanced concepts of inertial navigation. © *2004 MAIK "Nauka/Interperiodica".*

1. It is known [1] that absolute linear acceleration can be represented in the form $w = g + f$, where g is termed gravitational acceleration, free-motion (free-fall) acceleration, or gravitational field strength, and f is the apparent acceleration or the (specific) resultant of nongravitational forces.

In modern inertial navigational systems, linear acceleration is measured by either mechanical meters [1] or interferometers [2]. While being related to optical devices, the latter are functionally similar to the former, since interferometers, like mechanical meters, also measure apparent acceleration (f). This follows from the principle of their operation: the action of inertial mass on the material of the optical channel.

In this study, which pursues our discussion of measurement interpretation [3], we estimate the functionality of an, in a sense, purely optical accelerometer built around the following scheme. Let a source of optical radiation with frequency ν be placed at end A of linear segment AB, which has a given length L. Also, let reflectors be positioned at both ends (A and B), so that multiple reflections of the radiation emitted from source A are possible. Generally speaking, the passage of the radiation in the forward (AB) and reverse (BA) directions takes place in various optical media with refractive indices n_1 and n_2, respectively. After N passes, which constitute one measuring cycle, the observer sees the frequency ν_N measured.

We will deal with inertial coordinate system K and device-related system $\tilde{K}$ with origin o in the middle of section AB. The latter system moves along a straight line with absolute acceleration w, which is collinear to section AB. The following events should be treated by the observer, who is in system $\tilde{K}$.

2. Since the accelerated motion of system $\tilde{K}$ is of additive character, it makes sense to consider first the case $f \equiv 0$ and $g = \text{const}$. Taking into account that the gravitational field is present in both systems K and $\tilde{K}$ and also that inertial force J_g is present in the latter system (so that $J_g + g = 0$, the well-known zero-gravity phenomenon), we may argue that the radiation frequency measured in $\tilde{K}$ does not change. Otherwise, the effect of gravitational (red) shift supported experimentally [4, 5] comes into question.

In system $\tilde{K}$, moving with acceleration $w = g + f$ ($f \neq 0$, $g = \text{const}$), an inertial force $J_f = -f$ arises, which is not balanced by the external gravitational field. In accordance with the equivalence principle [4], this force can be identified with the strength of the gravitational field that exists only in $\tilde{K}$. This fact makes it possible to observe the corresponding "gravitational" frequency shift directly in $\tilde{K}$. Also, this fact (combined with the above consideration) means that the direct optical method considered in this study measures (in $\tilde{K}$) only apparent acceleration f (at $g = \text{const}$).

3. In the general case of the nonuniform gravitational field ($g \neq \text{const}$), we will determine the intrinsic time in $\tilde{K}$ in terms of the well-known interpretation of the metric tensor (for weak fields, the tensor is considered in [4, 5]). At the ends of segment AB in an ith optical channel ($i = 1, 2$), the tensor component g_{00}, which corresponds to time in four-dimensional space, has the following form:

$$g_{00}^{j,i} = 1 + 2\varphi_j/c_i^2, \quad i = 1, 2; \quad j = A, B \text{ (or } j = 1, 2), \tag{1}$$

where $c_i = c/n_i$, c is the velocity of light in free space, and φ_j ($j = A, B$) are the force potentials at ends A and B.

Quantities φ_j ($j = A, B$) can be represented as the sums of two components related to the gravitational forces ($\varphi_{j,g}$) and inertial forces ($\varphi_{j,J}$), so that $\varphi_j = \varphi_{j,g} + \varphi_{j,J}$ ($j = A, B$). Denoting the axis of system K that is collinear to AB by $0x$, we expand the partial potentials in

1063-7842/04/4909-1247$26.00 © 2004 MAIK "Nauka/Interperiodica"

the vicinity of the point o:

$$\varphi_{j,g} = \varphi_0 + (-1)^j \frac{\partial \varphi_g}{\partial x}\frac{L}{2} + \frac{1}{2}\frac{\partial^2 \varphi_g}{\partial x^2}\left(\frac{L}{2}\right)^2 + (-1)^j \frac{1}{6}\frac{\partial^3 \varphi_g}{\partial x^3}\left(\frac{L}{2}\right)^3 + \ldots,$$
$$\varphi_{j,J} = (-1)^j f\frac{L}{2} + (-1)^{(j-1)}\frac{\partial \varphi_g}{\partial x}\frac{L}{2}; \quad (2)$$
$$j = 1, 2 \text{ (or } j = A, B).$$

Here, φ_g is the gravitational potential and φ_0 is its value at point o. All partial derivatives are also taken at this point.

If the relationship between the intrinsic time τ of the points in $\tilde{K}$ and time t in inertial system K upon completion of the measuring cycle is taken into account, $d\tau^{j,i} = (g_{00}^{j,i})^{1/2}dt$ ($j = A, B$; $j = 1, 2$), the frequency measured takes the form

$$\nu_N = \nu[g_{00}^{B,1}g_{00}^{A,2}/(g_{00}^{A,1}g_{00}^{B,2})]^{N/2} \approx \nu + \delta\nu_N.$$

Here, the value of red (or, possibly, blue) shift $\delta\nu_N$ is the information-bearing component, which, in view of (2), can be expressed in the form

$$\delta\nu_N = \frac{\nu N(n_1^2 + n_2^2)}{c^2}(fL + \Delta),$$
$$\Delta = -\sum_{n=1}^{\infty}\frac{2}{(2n-1)!}\frac{\partial^{2n-1}\varphi_g}{\partial x^{2n-1}}\left(\frac{L}{2}\right)^{2n-1}. \quad (3)$$

From Eq. (3), we get

$$f + \Delta/L = c^2(\nu_N - \nu)/(\nu NL(n_1^2 - n_2^2)). \quad (4)$$

The order of magnitude of the values involved in (3) and (4) and the measurement accuracy required can be estimated from the following example. If φ_g is identified with the exterior field of terrestrial gravitation, which is assumed to be central, and the motion is considered along the central straight line, then $\Delta L \approx gL^2/8r^2$ (r is the distance to the center of the Earth). Setting L = 1 m, $n_1 = 1$, $n_2 = 2$, $\nu = 3 \times 10^{14}$ Hz, $N = 10^8$, $f \in [10^{-4}; 10]$ m/s^2, and $r = 6.4 \times 10^6$ m, we arrive at $\delta\nu_N \in [1; 10]$ Hz and $\Delta/L \approx 3 \times 10^{-14}$ m/s^2.

As is seen from the example, the fact that Eq. (4) includes quantity Δ/L, which contains information on the gravitational field, places more stringent requirements on the accuracy of measurement. Note that the optical accelerometer under study can also be employed as a gravimeter. In this case, the specific bearing reaction of the device's support, which is equal to the local gravitational field strength, plays the role of a specific nongravitational force (i.e., f).

4. Thus, it is shown that optical accelerometers, which are based on the effect of relativistic frequency shift, provide information on both the apparent acceleration (f) and the gravitational field (Δ). In the context of existing measurement concepts and inertial navigation practice, these devices can be viewed as meters of apparent acceleration in the case of a slightly nonuniform field (such as the exterior field of terrestrial gravitation); in other words, they are functionally similar to the mechanical and interferometric accelerometers currently available.

It is hoped that measurement techniques will advance to a level allowing one to record microacceleration on the order of (10^{-14}–10^{-16}) m/s^2. This will substantially raise the information content of both measurement components, which, in turn, will have a profound effect on the further development of the inertial navigation method [6].

REFERENCES

1. A. Yu. Ishlinskiĭ, *Classical Mechanics and Inertial Forces* (Nauka, Moscow, 1987) [in Russian].
2. V. I. Busurin and Yu. R. Nosov, *Optical-Fiber Sensors: Physical Grounds, Calculations, and Applications* (Énergoatomizdat, Moscow, 1990) [in Russian].
3. A. S. Devyatisil'nyĭ, Zh. Tekh. Fiz. **73** (9), 130 (2003) [Tech. Phys. **48**, 1209 (2003)].
4. W. Pauli, *The Theory of Relativity* (Pergamon, Oxford, 1958; Nauka, Moscow, 1991).
5. W. Burke, *Space–Time, Geometry, Cosmology* (Univ. Science Books, California, 1980; Mir, Moscow, 1985).
6. A. S. Devyatisil'nyĭ, Zh. Tekh. Fiz. **74** (5), 134 (2004) [Tech. Phys. **49**, 651 (2004)].

Translated by A. Sidorova