Technical Physics, Vol. 50, No. 2, 2005, pp. 141–145. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 1–5.
Original Russian Text Copyright © 2005 by Shitov, Moskalev.

THEORETICAL AND MATHEMATICAL PHYSICS

Modification of the Voss Algorithm for Simulation of the Internal Structure of a Porous Medium

V. V. Shitov and P. V. Moskalev
International Institute of Computer Technologies, Voronezh, 394026 Russia
e-mail: moskaleff@mail.ru

Received February 25, 2004

Abstract—The Voss random addition algorithm is modified for simulation of a stochastic fractal function that describes the distribution of the capillary system characteristics in a locally inhomogeneous porous medium with regard to porosimetry data. The influence of the dispersion function shape and the density of pores on the distribution of the capillary equivalent diameter in a given domain is analyzed. *© 2005 Pleiades Publishing, Inc.*

A wide range of structural and thermal physical properties, the ease of producing intricate parts, the developed internal surface, and intense heat exchange all make it possible to apply powder capillary–porous materials in various environments.

To work out physically correct methods for thermal and hydraulic simulation of devices containing porous elements, it is necessary to have reliable information on the internal structure of and mass and energy transfer in penetrable materials.

An important stage in solving these problems is development of an effective algorithm to simulate the distribution of the equivalent diameter (as well as any other parameter of a porous medium) from the generalized characteristics of the medium in a given region.

Note that some of the researchers who examined the structure of capillary–porous materials considered the characteristic dimensions of a porous medium as a whole [1, 2]. However, when studying the dynamics of fluids in the internal channels of voids, one has to take into account the size and velocity profile distributions over the cross section of the channel and also a decrease in the effective cross-sectional area of the channel due to local flow separation from the walls. Therefore, it seems to be worthwhile using the notion of the equivalent diameter (by analogy with noncircular channels), which has gained acceptance in applied hydrodynamics [3].

At present, there are strong grounds to believe that most locally inhomogeneous penetrable structures have the properties of stochastic self-similarity in a certain range of characteristic scales [4].

One of the most efficient procedures for constructing a spatial distribution with fractal characteristics is a generalization of the random addition algorithm, which was first proposed by Voss at a Conference on Fundamental Algorithms in Computer Graphics (Berlin, 1985) [5].

The basic version of the Voss algorithm represents a recursive sequence of the additions of the initial values of some pseudorandom function d_{i-1} with increments Δd_i that obey an unbiased normal distribution law with given rms deviation σ_i. A change in the number of points is specified by the partition coefficient

$$r = \Delta l_i / \Delta l_{i-1}, \tag{1}$$

where Δl_i and Δl_{i-1} are the increments of an independent variable for two sequential generations of the fractal curve. For constructing a one-dimensional Voss curve, we have to require that the dispersion of the increments of a dependent variable meet the relationship

$$\sigma^2(d_i) = r^{2H}\sigma^2(d_{i-1}), \tag{2}$$

where $\sigma^2(d_i)$ is the dispersion of the increments of the dependent variable for an ith generation of the random addition sequence and H is the Hurst exponent (which, in the general case, belongs to the open interval $H \in (0; 1)$).

For this procedure to be algorithmically realized for each element from the range of a current-generation function, it is convenient to use a recursive representation in the form of a normally distributed pseudorandom number with a mathematical expectation that is equal to the value of the previous-generation function and with a dispersion that is determined from (2).

The number of points for determination of a fractal function from generation to generation also has a recursive representation where partition coefficient (1) has the form

$$N(d_i) = rN(d_{i-1}). \tag{3}$$

Let us complete the set of Eqs. (2) and (3) by an expression for a dimensionless spatial resolution $\delta(d_i)$

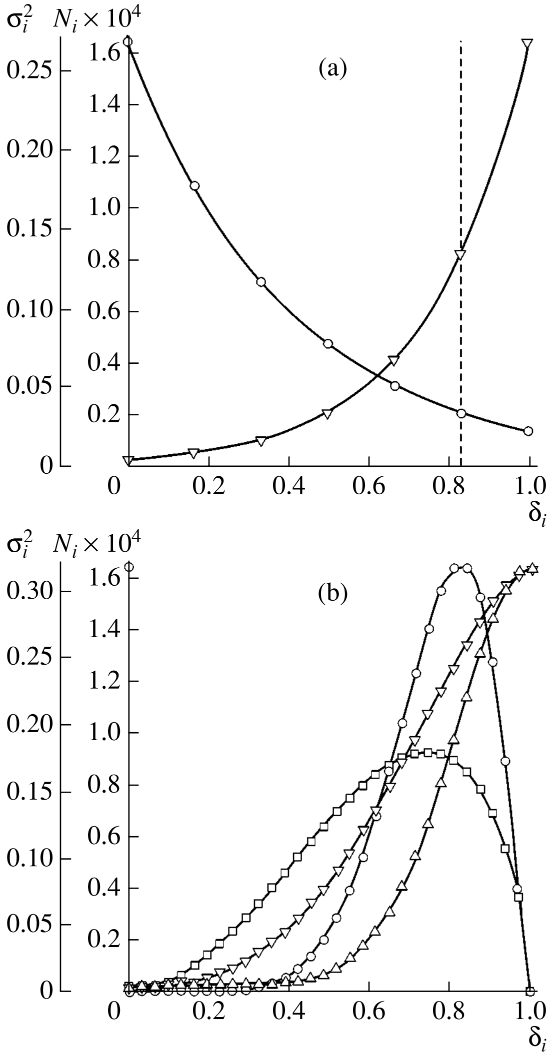


Fig. 1. σ_i^2 and N_i vs. current spatial resolution δ_i obtained by using the (a) standard and (b) modified Voss random addition algorithm. (–○–) and (–□–) refer to $\sigma^2(d_i)$ and (–△–) and (–▽–), to $N(d_i)$. The basic parameters for these distributions are as follows. (a) $\delta(d_a) = 1.000$ and $\delta(d_a)' = 0.833$; (–○–) $H = 0.600$ and $\sigma^2(d_b) = 0.268$; and (–△–) $r = 0.500$, $N(d_a) = 1.638 \times 10^4$, and $N(d_a)' = 0.410 \times 10^4$. (b) (–○–) $\alpha_\sigma = 8.0$, $\beta_\sigma = 2.5$, and $(\sigma^2)_{\max} = 0.320$; (–□–) $\alpha_\sigma = 3.0$, $\beta_\sigma = 1.7$, $(\sigma^2)_{\max} = 0.181$; (–△–) $\alpha_N = 8.0$, $\beta_N = 2.5$, and $N(d_a) = 1.638 \times 10^4$; and (–▽–) $\alpha_N = 3.0$, $\beta_N = 1.7$, and $N(d_a) = 1.638 \times 10^4$.

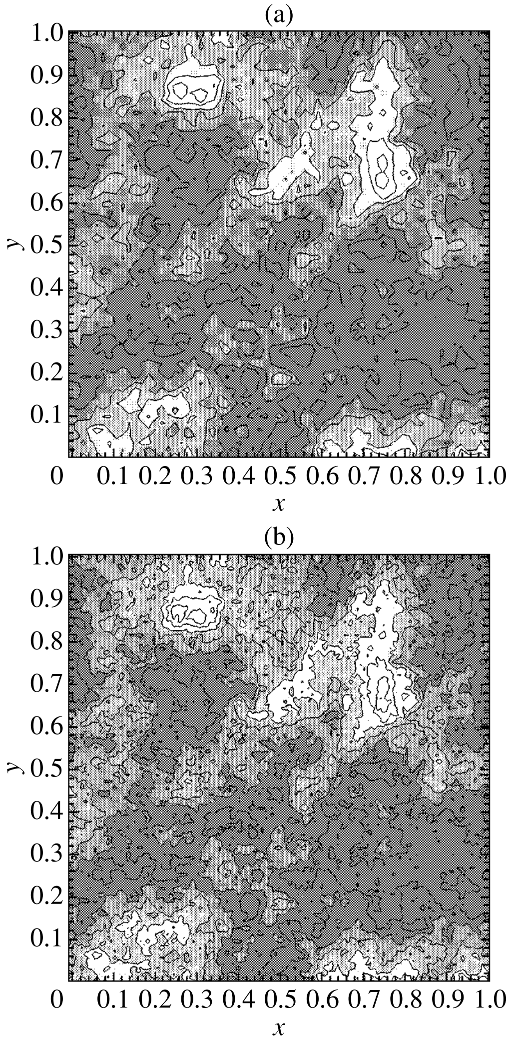


Fig. 2. Two realizations of the standard Voss random addition algorithm for stochastic fractal function $d(x, y)$ that describes the distribution of the equivalent hydraulic diameter of voids in a locally inhomogeneous porous medium. The basic parameters for these realizations are the following: (a) $\sigma^2(d_b) = 0.268$, $\delta(d_a) = 0.833$, and $N(d_a) = 0.410 \times 10^4$; (b) $\sigma^2(d_b) = 0.268$, $\delta(d_a) = 1.000$, and $N(d_a) = 1.638 \times 10^4$. For both realizations, $r = 0.500$ and $H = 0.600$.

normalized to the amplitude values of variable d_i:

$$\delta(d_i) = (d_b - d_i)/(d_b - d_a), \qquad (4)$$

where d_a and d_b are the minimal and maximal values in the range of independent variable d_i.

The independent solution of Eqs. (2) and (3) in view of spatial resolution (4) (Fig. 1a) shows that the standard realization of the Voss algorithm implies a monotonic increase in the number of points, $N(d_i)$, in each next generation of stochastic fractal function d_i with a similar monotonic decrease in the dispersion $\sigma^2(d_i)$ of its additional amplitude values Δd_i.

In addition, we note that this random addition algorithm admits a trivial generalization for the case of constructing a fractal function of two or more independent variables.

The realization of such a function for two values of maximal spatial resolution $\delta(d_a)$, with all other parameters unchanged, is exemplified in Fig. 2a ($\delta(d_a) = 0.833$) and 2b ($\delta(d_a) = 1.000$). The minimum value of the spatial resolution was kept constant ($\delta(d_b)$ = const) in all the realizations.

As applied to simulation of penetrable powder porous materials with a given grain composition, this means that, in the former case, the finest grains are excluded from the initial material.

The initial state of the function generating a pseudorandom number sequence, which governs the local distribution of increments Δd_i of the desired function, was set to be identical in both cases.

However, when using this algorithm in applications where the spatial structure of locally inhomogeneous capillary networks is to be simulated, one has to take into account the distributions of $\sigma^2(d_i)$ and $N(d_i)$ at different spatial resolution levels in real porous materials.

Experimental data on the structural characteristics of porous materials used in machine building [2, 6] indicate that the behavior of the dispersion function and the equivalent capillary diameter density function differ from that shown in Fig. 1a.

A typical distribution of the equivalent pore diameter in powder porous materials is shown in Fig. 3. The materials studied were BrOF10-1 tin–phosphor cast bronze (State Standard 1.90054-72) (Fig. 3a) and 12Kh18N10T stainless steel (State Standard 5632-72) (Fig. 3b). The experimental curves were obtained using the method of invasive mercury porosimetry at the Bauman State Technical University [6].

According to the test results, the voidage versus the equivalent pore diameter for all fractions no finer than a current fraction is given by

$$\Pi(d_i) = \frac{\Sigma_{k-i}^{n}\Delta V_k}{V_0}, \qquad (5)$$

where ΔV_k is a voidage increment due to the current fraction d_i of the porous material when the mercury pressure increases and V_0 is the total volume of the porous sample studied.

In Fig. 3, curve –▽– shows the integrated dependence of the total number of pores, $N(d_i)$, on their equivalent diameter for all fractions no finer than the current fraction. This curve is constructed from the experimental dependence $\Pi(d_i)$ under the following assumptions: (i) the pore size within a fraction is constant and equal to the mean equivalent diameter for the fraction, (ii) the pores are spherical and may intersect each other, and (iii) the fraction of the volume that is lost when two pores intersect is equal to, or larger than, half a smallest pore.

In this case, the lower estimate of the number of pores for each fraction depends on voidage ΔV_i due to this fraction,

$$N(d_i) = 2\frac{\Delta V_i \times 6}{\pi d_i^3}. \qquad (6)$$

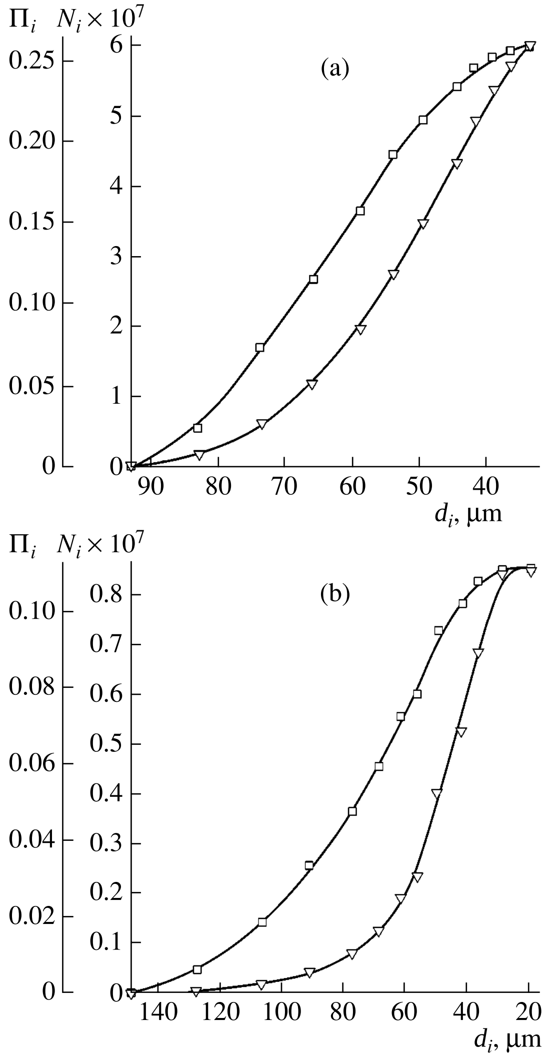


Fig. 3. Integral distributions of (□) voidage Π_i and (▽) number of pores N_i for the main fractions vs. mean equivalent diameter d_i. The basic parameters for the samples under study are the following: (a) porous bronze, d_0 = 25 mm, h_0 = 20 mm, Π_0 = 0.260, and $N_0 = 6.067 \times 10^7$; (b) porous stainless steel, d_0 = 20 mm, h_0 = 10 mm, Π_0 = 0.111, and $N_0 = 0.858 \times 10^7$.

Analysis of distribution functions $N(d_i)$ obtained by jointly solving Eqs. (5) and (6) indicates that the experimental data are best fitted by a family of integral beta-distribution functions [7],

$$N(d_i) = \frac{1}{\mathrm{B}(\alpha, \beta)}\int_{d_a}^{d_i}(d_i)^{\alpha-1}(1-d_i)^{\beta-1}d(d_i). \qquad (7)$$

Here, α and β are the distribution parameters and $\mathrm{B}(\alpha, \beta)$ is the beta function defined by the integral relationship

$$\mathrm{B}(\alpha, \beta) = \int_0^1 (t)^{\alpha-1}(1-t)^{\beta-1}dt, \qquad (8)$$

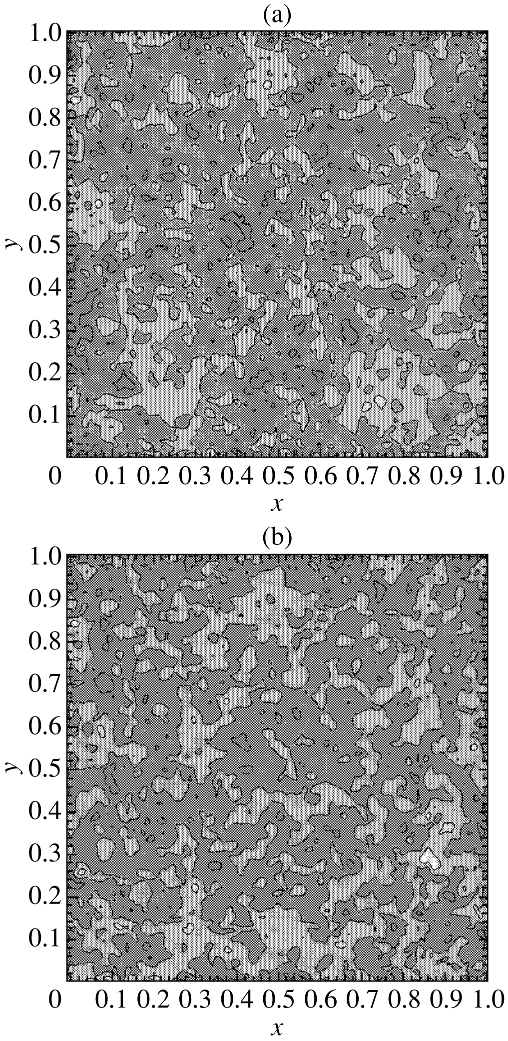


Fig. 4. Two realizations of the modified Voss random addition algorithm for stochastic fractal function $d(x, y)$ that describes the distribution of the equivalent hydraulic diameter of voids in a locally inhomogeneous porous medium. The distributions of the basic parameters for these realizations over spatial resolution levels are shown in Fig. 1b. Structurally, the realizations correspond to the samples (Fig. 3) of the (a) porous bronze ($\alpha_\sigma = \alpha_N = 3.0$ and $\beta_\sigma = \beta_N = 1.7$) and (b) porous stainless steel ($\alpha_\sigma = \alpha_N = 8.0$ and $\beta_\sigma = \beta_N = 2.5$).

where

$$t = \frac{d_x - d_a}{d_b - d_a}$$

is the variable of integration.

The approximations of the number of pores versus equivalent diameter dependence, $N(d_i)$, using Eqs. (7) and (8) are exemplified in Fig. 1b.

When simulating the dispersion distribution over spatial resolution levels, we used the relationship

$$\sigma^2(d_i) = \frac{1}{\mathrm{B}(\alpha, \beta)}(d_i)^{\alpha-1}(1-d_i)^{\beta-1}I(d_i), \qquad (9)$$

for the differential beta function. Here, $I(d_i)$ is the indicator of the event $\{d_i < d_x\}$, which has the form of the piecewise constant function

$$I(d_i) = \begin{cases} 1, & d_i < d_x \\ 0, & d_i \geq d_x. \end{cases}$$

In all the cases under consideration, the coefficients of the $\sigma^2(d_i)$ distribution were taken to be identical to those of the $N(d_i)$ distribution, $\alpha_\sigma = \alpha_N$ and $\beta_\sigma = \beta_N$. With such a choice, function σ_i^2 is asymmetric and peaks in the range where the gradient of function N_i reaches a maximum (Fig. 1b). Such behavior satisfactorily agrees with the experimental differential distributions of pores over their equivalent sizes in the materials under study [6].

Two realizations of the modified Voss algorithm using distribution functions N_i and σ_i^2 in the form of (7) and (9) are shown in Fig. 4. Structurally, these realizations correspond to the parameters of the samples, namely, the porous bronze (Fig. 4a) and the porous stainless steel (Fig. 4b). The only discrepancy is in the total number N_0 of equivalent pores, which form the capillary structure of the samples.

The total number of pores in real porous materials is extremely large. Subject to the above assumptions, it is determined from porosimetry data as the integral sum of the number of pores over all fractions. The lower estimates are as follows: for a cylindrical sample of the porous bronze (d_0 = 25 mm, h_0 = 20 mm, and voidage $\Pi_0 = 0.260$), the number of pores is $N_0 = 6.067 \times 10^7$; for a sample of the porous stainless steel (d_0 = 20 mm, h_0 = 10 mm, and $\Pi_0 = 0.111$), we have $N_0 = 0.858 \times 10^7$.

When simulating the capillary system of a porous body with the Voss algorithm, we restricted the dimension of the computational grid to $N_0' = (128)^2 = 1.638 \times 10^4$ in order to cut the machine time.

This restriction means a decrease in the linear dimensions of the region being simulated; in the two-dimensional case, this decrease is proportional to the ratio

$$\mu_0 = \frac{l_0'}{l_0} = \sqrt{\frac{N_0'}{N_0}}, \qquad (10)$$

where l_0' and l_0 are the characteristic linear dimensions of the virtual (simulated) and real regions, respectively.

Solving Eq. (10) for unknown variable l_0' yields lower estimates of the characteristic linear dimensions of the regions shown in Fig. 4: for the porous bronze simulant with $l_0 = d_0 = 25$ mm, we obtain $l'_0 = 0.411$ mm; for the porous stainless steel simulant at $l_0 = d_0 = 20$ mm, we get $l'_0 = 0.874$ mm.

Thus, when compared with the standard version, the modified Voss algorithm, which takes into account the behavior of the structural indices in real capillary–porous materials on various spatial scales, substantially changes the structure of the resulting capillary system. First, the small value and the slow increase of index σ_i^2 in the range of low spatial resolutions δ_i produce a more uniform stochastic distribution of equivalent diameters. This is because the portion of coarser fractions decreases and basic fluctuations of function $d(x, y)$ shift toward medium scales. Second, the sharp decline in index σ_i^2 in the range of high δ_i considerably smoothes stochastic function $d(x, y)$, because the finest fractions become minor contributors to the resulting equivalent diameter distribution.

CONCLUSIONS

The modified version of the Voss algorithm offers a high flexibility, which is very important in applied research. The beta functions applied to simulate the differential and integral distributions allow one to reproduce most practically important smooth functions in a limited interval. Moreover, these functions may be specified in tabular form (e.g., based on empirical data) without any restrictions and missing values may then be found by interpolation using reference points. No restrictions are also imposed on interpolation methods. It is shown that a porous structure under study can be reproduced from porosimetry data as fully as desired. Owing to its recursive structure, the modified algorithm generates stochastically self-similar realizations of a multidimensional fractal function with characteristics specified in a given domain of the space of independent variables. The aforesaid suggests that the procedure proposed in this work could be helpful in theoretical and applied studies of locally inhomogeneous porous structures. However, a relationship between the fractal characteristics of a function thus obtained and the parameters of the initial distributions has yet to be reliably established.

REFERENCES

1. Yu. V. Polezhaev and F. B. Yurevich, *Thermal Protection* (Énergiya, Moscow, 1976) [in Russian].
2. V. M. Polyaev, V. A. Maĭorov, and L. L. Vasil'ev, *Hydrodynamics and Heat Exchange in the Porous Elements of Aircrafts* (Mashinostroenie, Moscow, 1988) [in Russian].
3. I. E. Idel'chik, *Handbook on Hydrodynamic Drag* (Mashinostroenie, Moscow, 1975) [in Russian].
4. J. Feder, *Fractals* (Plenum, New York, 1988; Mir, Moscow, 1991).
5. R. F. Voss, *Fundamental Algorithms in Computer Graphics* (Springer-Verlag, Berlin, 1985), pp. 805–835.
6. S. V. Belov, *Porous Metals in Machine Building* (Mashinostroenie, Moscow, 1981) [in Russian].
7. M. Evans, N. Hastings, and J. B. Peacock, *Statistical Distributions*, 3rd ed. (Wiley, New York, 2000).

Translated by K. Shakhlevich

Technical Physics, Vol. 50, No. 2, 2005, pp. 146–151. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 6–12.
Original Russian Text Copyright © 2005 by Abramovich, Zavjalov, Zvenigorodsky, Ignat'ev, Magilin, Melnik, Ponomarev.

THEORETICAL AND MATHEMATICAL PHYSICS

Optimization of the Probe-Forming System for a Scanning Nuclear Microprobe Based on the ÉGP-10 Electrostatic Tandem Accelerator

S. N. Abramovich*, V. N. Zavjalov*, A. G. Zvenigorodsky*, I. G. Ignat'ev,**
D. V. Magilin, K. I. Melnik**, and A. G. Ponomarev****
** Institute of Nuclear and Radiation Physics, Russian Federal Nuclear Center,*
Research Institute of Experimental Physics (VNIIEF), Sarov, Nizhni Novgorod Oblast, 207190 Russia
*** Institute of Applied Physics, National Academy of Sciences of Ukraine,*
ul. Petropavlovskaya 58, Sumy, 40030 Ukraine
e-mail: ipfmail@ipfcentr.sumy.ua

Received March 31, 2004; in final form, June 30, 2004

Abstract—The probe-forming system of a nuclear scanning microprobe based on the parametric multiplets of quadrupole lenses is optimized. The optimization is aimed at creating an ion probe with energy of several MeV that produces a micrometer spot on the target at a current of ~100 pA. The influence of different geometric and physical parameters on the ion–optical properties of the probe-forming systems considered is determined. The optimization is carried out by varying the parameters specifying a given parametric multiplet, and its efficiency is found from a quality criterion that takes into account the beam current for given sizes of the spot and target. The beam parameters at the entrance to and at the exit from the ÉGP-10 electrostatic tandem accelerator (produced by the VNIIÉF) are involved in the optimizing calculations. These are the maximal energy, normalized brightness, transport conditions, and chromatic inhomogeneity of the beam (i.e., the energy straggling of beam particles). Allowance is also made for the parasitic components of the magnetic quadrupole lens field, which arise because of quadrupole symmetry breaking by technological and physical reasons. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

Electrostatic accelerators (EAs), which were initially developed for solving the problems of experimental nuclear physics, continue to be promising for various fields of science and technology. In the first place, this is because the beams used in EAs are highly monochromatic and their energy can be smoothly varied. An interesting application is an EA-based nuclear scanning microprobe (NSMP) intended for studying the structure and elemental composition of various, including biological, materials by such matured methods as particle-induced X-ray emission (PIXE), nuclear reaction analysis (NRA), Rutherford back scattering (RBS), and others. NSMPs have found wide utility in producing parts for micromachines [1] and in precision ion implantation [2]. The need for a probe using focused ion beams with energy of several MeV that produces a micrometer spot on the target has stimulated design of probe-forming systems (PFSs) capable of providing the desired parameters. Experience gained by the researchers engaged in electron microscopy studies, where focusing units generate axisymmetric fields, may help little in the case of an NSMP because of the high energy of the beam. In most advanced nuclear microprobes, a magnetic quadrupole lens serves as a basic focusing element. According to [3], the relativistic effects should be taken into consideration when the particle velocity v is such that $v/c > 0.2$, where c is the speed of light. For protons with energy of 14 MeV, $v/c \approx 0.17$. Therefore, the charged particle motion here is assumed to be nonrelativistic. The power of a magnetic quadrupole lens depends on the dimensionless excitation κ of a hyperbolic-shape pole piece: $\kappa = L_E\sqrt{W_2\sqrt{2q/mV}}$. Here, W_2 is the quadrupole component, which is governed by the field gradient in the direction transverse to the optical axis; L_E is the effective length of the lens field; m and q are the ion charge and mass, respectively; and V is the potential difference across the accelerating gap. Unlike axisymmetric lenses, the power of a quadrupole lens can be raised by increasing not only the field gradient but also the effective length of the lens field. Hence, there are no major restrictions on the beam energy. Since quadrupoles can focus in only one transverse direction, the stigmatic focusing of the beam on the target is provided with multiplet systems consisting of several lenses. The arrangement of quadrupoles in a multiplet is shown in Fig. 1.

The PFS itself consists of object and angular collimators, focusing elements, and the target. Their position on the optical axis is specified by geometric parameters a, g, a_i, l, and $L_{E,(i)}$, where a is the object distance, i.e., the distance between the object and angular collimators; g is the working distance, i.e., the distance between the exit boundary of the effective field of the last lens to the target plane; a_i specifies the position of an ith lens; l is the length of the system; and $L_{E,(i)}$ is the effective length of the field of the ith lens.

Each of the lenses is fed by a current source that provides the distribution of the quadrupole component $W_{2,(i)}(z)$ along the optical axis (z axis). The local Cartesian coordinate system of each of the lenses is related to the optical axis z of the lens and its planes of quadrupole antisymmetry. The optical axes of the lenses are aligned with the optical axis Z of the laboratory coordinate system. The local coordinate systems of the lenses may either coincide with the laboratory system or be rotated through ±90° about the Z axis.

The PFS thus defined will be called the PFS based on the parametric multiplets of magnetic quadrupole lenses. Here, variable parameters may be the number N of lenses in the system and the number n of independent power suppliers; parameters l, a, a_i, g, and $L_{E,(i)}$ are defined above. Triplets ($N = 3$) and quadruplets ($N = 4$) of magnetic quadrupole lenses with two independent power suppliers ($n = 2$) are today well understood and extensively used. The triplet that has found the widest recognition and, in particular, is used in the microprobe at the Oxford laboratory [4] is fed as follows: the first and second lenses are fed by one source, and the third one is fed independently by another source. The local coordinate system of the second lens is rotated through ±90° about the Z axis of the laboratory coordinate system. In the PFS known by the name *Russian quadruplet* ($N = 4$) [5], the first and fourth lenses are fed from one source and the second and third ones, from another independent source. The local coordinate systems of the second and fourth lenses are rotated through ±90° about the Z axis of the laboratory coordinates. The systems with $N > 4$ and $n > 2$ are yet little understood. The quintuplet ($N = 5$, $n = 2$) of magnetic quadrupole lenses is applied in microprobes in Melbourne [6] and Sydney [7].

When optimizing the PFS, we consider the beam parameters (the energy E_0, normalized brightness b, and ion energy straggling $\delta_E = \Delta E/E_0$) provided by the ÉGP-10 VNIIÉF accelerator. The dimensions of the lenses and their fields take into account the magnetic properties of the material. In the process of PFS operating parameter optimization, the geometric parameters of the system are varied. The analysis also includes the nonlinear effects associated with both intrinsic and parasitic aberrations due to technological imperfections. The optimal PFS meets a quality criterion that is defined as the current provided by the system when the beam is focused into a spot of a given size.

ION–OPTICAL PROPERTIES OF PROBE-FORMING SYSTEMS

By the ion–optical properties of a PFS, we mean the demagnification and aberration coefficients. Aberrations include first-order chromatic aberrations; third-order intrinsic aberrations; and second- and third-order parasitic aberrations, which are due to the sextupole and octupole parasitic components of the lens field. The parasitic components arise when technology-related inaccuracies in fabricating the magnetic circuits and coils break the quadrupole symmetry of the field.

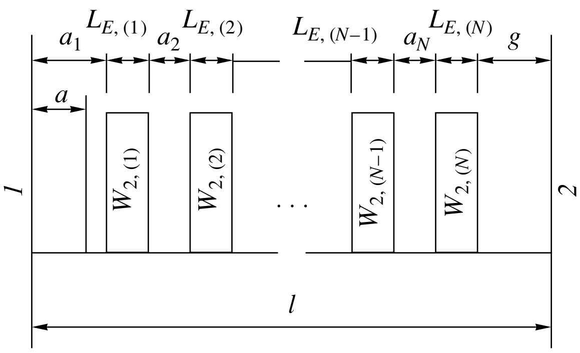


Fig. 1. Arrangement of quadrupole lenses in a multiplet. (*1*) Object and (*2*) target.

Field gradient $W_{2,(i)}(z)$ in each of the lenses must keep the stigmatism of the system for given parameters of a parametric multiplet. The aberrations were calculated by the matrizant method [8]. In terms of matrix methods, the aberrations of the entire system are found by calculating those of each active element. Our approach to calculating the aberrations of a magnetic quadrupole lens with regard to contributions from the parasitic components of the lens field [9] is based on the axial field model. It is assumed that there is a rectilinear optical axis in the lens along which the field vanishes, $|\mathbf{B}| = 0$. This approximation implies that the particles do not change the direction of motion at the entrance to and the exit from the lens. In the local Cartesian system (x, y, z) the field generated by a magnetic quadrupole lens in the pole gaps is described by the scalar magnetic potential

$$\begin{aligned} w(x, y, z) \;=\; & 2W_2(z)xy + U_3(z)x^3 + 3W_3(z)x^2y \\ & - 3U_3(z)xy^2 - W_3(z)y^3 + U_4(z)x^4 \\ & + (4W_4(z) - W_2''(z)/6)x^3y - 6U_4(z)x^2y^2 \\ & - (4W_4(z) + W_2''(z)/6)xy^3 + U_4(z)y^4 + \ldots, \end{aligned} \tag{1}$$

where $W_2(z)$ is the major quadrupole component, $W_3(z)$ and $U_3(z)$ are the major and skew parts of the sextupole parasitic component, and $W_4(z)$ and $U_4(z)$ are the major and skew parts of the octupole parasitic component.

The coordinate system (x, y, z) that is related to the optical axis of the lens and the planes of antisymmetry of major quadrupole component $W_2(z)$ is termed the intrinsic coordinate system. Accordingly, components $W_3(z)$, $U_3(z)$, $W_4(z)$, and $U_4(z)$ are called the components of the intrinsic parasitics of a magnetic quadrupole lens. Expression (1) is obviously a superposition of the fields of $2n$-pole multipoles ($n = 2, 3, 4, \ldots$) where each of the multipole fields has its own local coordinate system coincident with its optical axis and planes of antisymmetry. Noncoincidence of this local coordinate system with the planes of antisymmetry because of rotation about the optical axis causes extra skew component $U_n(z)$ to appear. Since the parasitic multipole components are of random character, their local coordinate systems are freely oriented, i.e., are not related to the local coordinate system of the quadrupole component. Hence, they have both a major, $W_n(z)$, and skew, $U_n(z)$, component at $n > 2$. Thus, any minor rigid transformation of the lens is bound to redistribute the multipole components and/or to contribute to the parasitics of the dipole component and/or to the parasitics of the skew quadrupole component. The allowable level of the parasitics due to technological inaccuracies in lens fabrication, as well as field quadrupole symmetry breaking because of variously fed polar pieces, is considered in [9, 10].

The matrizant method maps the phase space (x, x', y, y') where the position and direction of motion of a particle in a magnetic field are described onto the space of phase moments. To calculate the mapping matrix for our magnetic quadrupole lens (whose field is described by scalar magnetic potential (1)), we represent the space of phase moments by the vector

$$\begin{gathered}\mathbf{X} = (x, x', y, y', x\delta, x'\delta, y\delta, y'\delta, x^2, xx', x'^2, y^2, yy',\\ y'^2, xy, x'y, xy', x'y', x^3, x^2x', xx'^2, x'^3, xy^2, xyy',\\ xy'^2, x'y^2, x'yy', x'y'^2, y^3, y^2y', yy'^2, y'^3, yx^2, yxx',\\ yx'^2, y'x^2, y'xx', y'x^2)^T.\end{gathered} \quad (2)$$

Here, x and y are the deviations of the beam particles from the optical axis, x' and y' are the angles between the velocity vector projections onto planes $x0z$ and $y0z$ and the z axis (these angles specify the direction of particle motion), and $\delta = (p - p_0)/p_0$ is the relative deviation of the particle momentum p from the mean value p_0.

In the matrizant method, the evolution of vector **X** along the z axis is given by

$$X(z) = \mathbf{R}(z/z_0)X(z_0). \quad (3)$$

Here, $\mathbf{R}(z/z_0)$ is the matrizant (the matrix mapping the coordinates of the 38×38-dimensional phase moment space from plane z_0 onto a plane with coordinate z) where the first and third rows contain a complete set of the linear properties of the magnetic quadrupole lens and aberrations (including chromatic aberrations, second- and third-order aberrations due to the sextupole and octupole components, and all intrinsic geometric aberrations of the third order). For the rectangular distribution of the field components $W_2(z)$, $W_i(z)$, and $U_i(z)$ ($i = 3, 4$), an analytic representation of the matrizant was obtained.

Based on the aforesaid, the program PROBFORM was developed, which makes it possible to determine the ion–optical properties of PFSs from the parametric multiplets of magnetic quadrupole lenses. The program was carefully tested on analytical models. The test results were compared with those of other authors and checked experimentally [11].

QUALITY CRITERION FOR A PROBE-FORMING SYSTEM

Since, in our work, the ion–optical properties of a PFS imply the demagnification and aberration coefficients, PFS optimization is aimed at designing systems with high demagnification and low aberration coefficients, since a high demagnification decreases the spot on the target, while aberrations, conversely, increase. Unfortunately, high demagnifications considerably enhance aberrations in all PFSs. Therefore, PFS optimization means finding a tradeoff between these coefficients for a particular system. By varying the parameters of parametric PFSs, one usually seeks for a system where the tradeoff between these coefficients maximizes the current density in the spot on the target. The current density in the spot is therefore taken as the quality criterion for PFSs. In other words, those systems are preferable that provide a maximal current at the target for a given size of the beam. Such a criterion meets the requirements of the chemical or physical analysis techniques that extract information from interaction events between beam particles and target atoms. Clearly, the rate of such events is in direct proportion to the number of particles incident on the spot per unit time, but the spot itself must be minimized in order to adequately examine the element distribution in the material. Therefore, the current density on the spot is limited from above in a number of applications. In this case, the minimal size of the spot on the target at a fixed current is taken as the quality criterion [12, 13]. It is known that $I \approx b\varepsilon E$, where I is the current and b, ε, and E are the normalized brightness, emittance, and energy of the beam, respectively. Since the normalized brightness characterizes the ion source and beam transport system and so depends on a specific accelerator design, a functional relationship $d = \hat{d}^*(\varepsilon)$ (where d is the minimal beam size at given emittance ε) should be known for each general-purpose PFS. From physical considerations, it follows that function $\hat{d}^*(\varepsilon)$ monotonically increases (d grows with ε). Then, the inverse dependence $\varepsilon = \hat{\varepsilon}^*(d)$ can be constructed by seeking for the maximal emittance of the beam with its size on the target fixed. A method based on this principle was suggested in [14]. This method of finding the maximal beam emittance is realized in the MaxBEmit numerical code and allows one to determine the maximum emittance provided by a PFS when its beam is focused into a spot of a given size.

SELECTION OF OPTIMAL PARAMETERS FOR A PROBE-FORMING SYSTEM

We investigated the effect of various geometrical and physical parameters on the ion–optical properties of the PFS based on the parametric multiplets of magnetic quadrupole lenses, using the quality criterion suggested above. With this criterion, one can choose the optimal PFS design for the ÉGP VNIIÉF electrostatic accelerator. The choice relies upon the following reasoning. It was shown [9, 15] that the Russian quadruplet is of higher performance than the Oxford triplet and the increase in the number of the lenses to five does not cause a significant increase in the current density on the target. Therefore, the PFS chosen uses the multiplet like the Russian quadruplet ($N = 4$, $n = 2$). Previous investigations [15] let us argue that, if the beam ion energy straggling in accelerators is $\delta_E \sim 10^{-3}$, the performance of such PFSs is virtually independent of their length in the range $2 < l < 9$ m and they cannot provide a submicron resolution at a current of ≈100 pA when used in PIXE, RBS, and NRA studies. This is because chromatic aberrations will necessitate smaller size collimators to be used in order to provide a submicron beam on the target; accordingly, the current will decrease. Therefore, it makes no sense to take long PFSs (l = 7–9 m), which offer a high demagnification ($D \approx 100$). The resolution is also limited by a relatively low normalized brightness of the beam, $b \approx$ 1 pA/(μm^2 mrad^2 MeV), which is typical of tandem accelerators with the standard ion source of duoplasmatron type. Finally, the arrangement of the microprobe channel should also be taken into consideration.

Based on the aforesaid, the length $l = 4$ m of the PFS was taken as the basic length of the NSMR. For the given basic length, it remains to determine the following parameters of the system: a, the object distance; g, the working distance; $L_{E,(1)}$ and $L_{E,(2)}$, the effective lengths of the lenses; r_a, the radius of the lens aperture; and W_i/W_2 and U_i/W_2 (i = 3, 4), the allowable relative levels of the field parasitic components in magnetic quadrupole lenses.

These parameters will specify the PFS that is optimal in terms of the maximum current for a given size of the spot on the target.

It was shown [9] that the allowable relative levels of the field parasitic components in the magnetic quadrupole lenses used in the PFSs based on the Oxford triplet and Russian quadruplet are $W_3/W_2 = U_3/W_2 =$ 0.000375 cm^{-1} and $W_4/W_2 = U_4/W_2 = 0.000538$ cm^{-2}, respectively. The aperture radius depends on the accuracy of positioning of the pole piece that can be provided in magnetic circuit fabrication. Our equipment (an electric spark machine with auxiliary facilities) provides a pole piece positioning accuracy of ±5 μm. Calculations show that the parasitics will fall into the allowable range in magnetic quadrupole lenses with an aperture radius of 6.5 mm at a pole piece positioning accuracy of < ±5 μm.

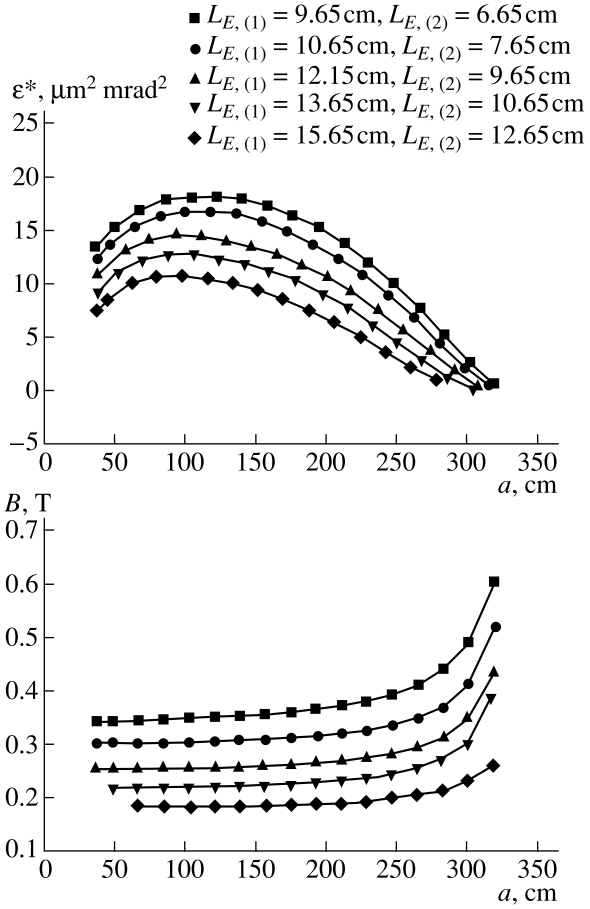


Fig. 2. Maximal emittance ε* and maximal magnetic induction B at the pole piece of the quadrupole lenses in PFSs vs. parameters a, $L_{E,(1)}$, and $L_{E,(2)}$.

Figure 2 shows the PFS characteristics versus parameters a, $L_{E,(1)}$, and $L_{E,(2)}$. Figure 1a plots the maximal emittance provided by the systems being considered when a beam of 14-MeV H$^+$ ions is focused into a square spot of size $d = 1.8$ μm at the given working distance $g = 20$ cm (parasitic aberrations are disregarded). In Fig. 1b, the vertical axis is magnetic induction B that provides the desired beam parameters on the target for given lengths of the lenses. The lengths of the lenses should be selected such that the maximal induction at the pole meets the requirement $B < 0.27$ T (in this case, we fall into the range of linearity of the pole piece material characteristics) and that the systems using such lenses provide a maximal emittance.

It follows from Fig. 2 that the systems with the effective lens lengths $L_{E,(1)}$ = 12.15 cm and $L_{E,(2)}$ = 9.65 cm offer the maximal emittance and satisfy the restrictions on the maximal magnetic induction at the pole piece. These lengths were taken as the basic dimensions of magnetic quadrupole lenses in the PFS.

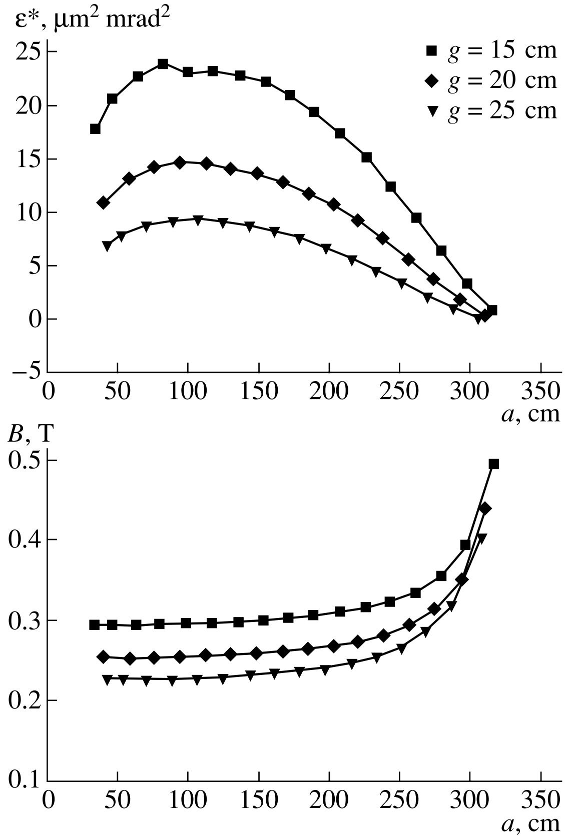


Fig. 3. Maximal emittance ε* and maximal magnetic induction B at the pole piece of the quadrupole lenses in PFSs vs. parameters a and g. The effective lens lengths are $L_{E,(1)}$ = 12.5 cm and $L_{E,(2)}$ = 9.65 cm.

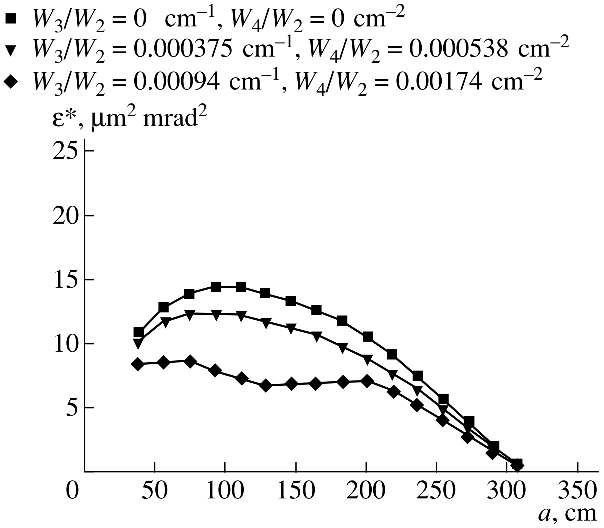


Fig. 4. Maximal emittance ε* in PFSs vs. object distance a and different levels of the field parasitic components in the magnetic quadrupole lenses.

Parameters of the basic probe-forming system (the Russian quadruplet of magnetic quadrupole lenses) for the EGP VNIIÉF accelerator

System length l, cm	400.0
Lens effective lengths $L_{E,(1)}$, cm	12.15
$L_{E,(2)}$, cm	9.65
Lens spacing in doublets $a_2 = a_4$, cm	3.0
Doublet spacing a_3, cm	129.0
Lens aperture radius r_a, cm	0.65
Working distance g, cm	20.0
Object distance $a = a_1$, cm	201.4
Maximal magnetic induction at pole B_{max}, T (E = 14 MeV)	0.269
Demagnifications: D_x	27.2(27.2)
D_y	27.2(27.2)
Chromatic aberrations, μm/mrad/%: $\langle x/x'\delta\rangle$	–607(–606)
$\langle y/y'\delta\rangle$	–125(–124)
Intrinsic spherical aberrations, μm/mrad³: $\langle x/x'^3\rangle$	807(840)
$\langle x/x'y'^2\rangle$	177(123)
$\langle y/y'^3\rangle$	16(19)
$\langle y/y'x'^2\rangle$	177(123)
Parasitic aberrations due to allowable sextupole components, μm/mrad²: $\langle x/x'^2\rangle$	–92
$\langle x/x'y'\rangle$	–17
$\langle x/y'^2\rangle$	3
$\langle y/x'^2\rangle$	–8
$\langle y/x'y'\rangle$	6
$\langle y/y'^2\rangle$	3
octupole components, μm/mrad³: $\langle x/x'^3\rangle$	370
$\langle x/x'y'^3\rangle$	–63
$\langle y/y'^3\rangle$	7
$\langle y/y'x'^2\rangle$	–63
Collimator size for focusing H⁺ beam, E = 14 MeV	
Square spot on target d = 1.8 μm, I = 100 pA	
Object collimator, $2r_x$ [μm]	27
$2r_y$ [μm]	25
Square spot on target d = 0.6 μm, I = 1.7 pA	
Object collimator, $2r_x$ [μm]	8
$2r_y$ [μm]	9

Figure 3 plots the same parameters of the systems as in Fig. 2 against g and a. As working distance g increases, so does the emittance, but magnetic induction B at the pole piece also grows, falling outside the allowable range (the range of linearity). Moreover, it is anticipated that the NSMP will employ an x–y ferromagnetic scanning unit with implicit pole pieces of active length 12 cm. Their design is similar to that described in [16]. To avoid lens aberrations, which cause beam parameter degradation at the edge of the raster, the scan unit is placed downstream of the last lens. For a maximal deflecting field of 0.02 T and working distance g = 20 cm, a necessary scan length of ±300 μm is provided. Analysis shows that, with such a scanning scheme, the increase in the beam diameter at the edge of the raster is no more than 2% and is related primarily to the chromatic inhomogeneity of the beam. Therefore, g = 20 cm is the lower limit of the working distance from both the physical and technological points of view.

Figure 4 shows the maximal emittance as a function of object distance a for lens lengths $L_{E,(1)} = 12.15$ cm and $L_{E,(2)} = 9.65$ cm, working distance g = 20 cm, and different levels of the field parasitic components. The system with a = 201 cm, which is the least sensitive to the parasitic components, is taken as the basic one.

The table summarizes the ion–optical properties, as well as the geometrical and physical parameters, of the basic PFS using the Russian quadruplet of magnetic quadrupole lenses. The parasitic aberrations are listed for the maximal permissible levels of the sextupole and octupole components of the lens fields. The aberration coefficients are designated starting from the map of the particle coordinates on the plane (x_o, y_o) of the object collimator onto the plane (x_t, y_t) of the target,

$$\alpha_t = \ldots + A^{[\alpha]}_{ijklm} x_o^i x_o'^j y_o^k y_o'^l \delta^m + \ldots,$$

where $A^{[\alpha]}_{ijklm} = \langle \alpha / x^i x'^j y^k y'^l \sigma^m \rangle$ are the aberration coefficients, $\alpha = \{x, y\}$; $i, j, k, l, m = 0, 1, 2, 3, \ldots$.

In the column showing the demagnifications, as well as the chromatic and intrinsic spherical aberrations, the related values calculated with the PRAM program are presented in the parentheses for comparison.

Thus, we optimized the PFS based on the parametric multiplets of magnetic quadrupole lenses. As a quality criterion of the system, the beam current at a given size of the spot on the target was taken. This value was the target function in the optimization problem. The geometric and physical parameters of the optimal PFS for the ÉGP VNIIÉF accelerator. The rated beam size on the target is 1.8 × 1.8 μm at a current of ≈100 pA, energy of 14 MeV, normalized brightness of ≈1 pA/(μm^2 mrad^2 MeV), and beam ion energy straggling of 10^{-3}.

REFERENCES

1. T. Osipowicz, J. A. van Kan, T. C. Sum, *et al.*, Nucl. Instrum. Methods Phys. Res. B **161–163**, 83 (2000).
2. J. Meijer, U. Weidenmuller, P. Baving, *et al.*, Nucl. Instrum. Methods Phys. Res. B **161–163**, 898 (2000).
3. M. Szilagyi, *Electron and Ion Optics* (Plenum, New York, 1988; Mir, Moscow, 1990).
4. D. N. Jamieson, G. W. Grime, and F. Watt, Nucl. Instrum. Methods Phys. Res. B **40–41**, 669 (1989).
5. A. D. Dymnikov and S. Ya. Yavor, Zh. Tekh. Fiz. **33**, 851 (1963) [Sov. Phys. Tech. Phys. **8**, 639 (1963)].
6. D. N. Jamieson, B. Rout, R. Szymanski, *et al.*, Nucl. Instrum. Methods Phys. Res. B **190**, 54 (2002).
7. C. G. Ryan and D. N. Jamieson, Nucl. Instrum. Methods Phys. Res. B **158**, 97 (1999).
8. A. Dymnikov and R. Hellborg, Nucl. Instrum. Methods Phys. Res. A **330**, 323 (1993).
9. A. G. Ponomarev, K. I. Melnik, V. I. Miroshnichenko, *et al.*, Nucl. Instrum. Methods Phys. Res. B **201**, 637 (2003).
10. L. A. Baranova and F. H. Read, Optik–Int. J. Light Electron Opt. **112**, 131 (2001).
11. V. A. Brazhnik, A. D. Dymnikov, D. N. Jamieson, *et al.*, Nucl. Instrum. Methods Phys. Res. B **104**, 92 (1995).
12. H. Wollnik and M. I. Yavor, Nucl. Instrum. Methods Phys. Res. B **158**, 113 (1999).
13. A. D. Dymnikov and G. Martinez, Nucl. Instrum. Methods Phys. Res. B **130**, 64 (1997).
14. A. G. Ponomarev, V. I. Miroshnichenko, and V. E. Storizhko, Nucl. Instrum. Methods Phys. Res. A **586**, 20 (2003).
15. V. A. Brazhnik, V. I. Miroshnichenko, A. G. Ponomarev, *et al.*, Nucl. Instrum. Methods Phys. Res. B **174**, 385 (2001).
16. V. Khomenko, S. Lebed, and S. Mordik, Nucl. Instrum. Methods Phys. Res. B **130**, 86 (1997).
17. M. Breese, D. Jamieson, and P. King, *Materials Analysis Using a Nuclear Microprobe* (Wiley, New York, 1996), p. 367.

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 152–157. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 13–18.
Original Russian Text Copyright © 2005 by Tereshchenko.

THEORETICAL
AND MATHEMATICAL PHYSICS

Emission Tomography of Radially Symmetric Objects and Exponential Abelian Transform

S. A. Tereshchenko
Moscow State Institute of Electronic Technology (Technical University),
Proezd 4806, 5, Zelenograd, Moscow, 124498 Russia
e-mail: tsa@miee.ru

Received March 9, 2004

Abstract—The properties of the exponential Abelian transform (EAT), which is defined as the exponential Radon transform (ERT) of a radially symmetric object just as the Abelian transform is the Radon transform of a radially symmetric object, are considered. A new approach to deriving the inverse EAT directly from the inverse ERT is suggested. The problems of numerical implementation of the EAT are discussed, including the problem of loss of information from deep-seated regions of the object, which is nonexistent in the case of the conventional Abel transform. The results obtained may be useful for reconstructing the spatial distribution of axisymmetric or spherically symmetric radiation sources. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

Computerized tomography is now a mature domain of science with its own scope of problems and solution methods [1–4]. Central in computerized tomography are the Radon transform [5], which arises in transmission tomography, for instance, in determining the spatial distribution of the radiation attenuation coefficient, and the exponential Radon transform [6–8], which is used in single-photon emission tomography when, for instance, the spatial distribution of radiation sources is sought. As applied to radially symmetric functions, the associated transformations are of particular importance. It turns out that the Radon transform of a radially symmetric function is the Abelian transform, which was well known before Radon and is today widely used in various domains of science [9]. At the same time, the exponential Radon transform of a radially symmetric function [10], which it is natural to call the exponential Abelian transform (EAT), has long been in disfavor, possibly because of the scarcity of relevant works.

The history of the integral Abelian transform (integral Abelian equation) dates back to 1823, when Niels Henrik Abel generalized the well-known problem of tautochrone. He obtained an integral equation of form

$$g(x) = \int_0^x \frac{f(t)}{\sqrt{x-t}}dt \tag{1}$$

and found its solution (the inverse transform)

$$f(x) = \frac{1}{\pi}\frac{d}{dx}\int_0^x \frac{g(t)}{\sqrt{x-t}}dt. \tag{2}$$

At present, the term *the Abelian equation (a generalized Abelian equation*) is usually referred to any equation of form

$$g(x) = \int_a^x \frac{f(t)}{(x-t)^\alpha}dt \tag{3}$$

(where $0 < \alpha < 1$ and $a > 0$) having a solution like

$$f(x) = \frac{\sin\alpha\pi}{\pi}\frac{d}{dx}\int_a^x \frac{g(t)}{(x-t)^{1-\alpha}}dt. \tag{4}$$

Integrating by parts, one obtains the inversion formula

$$f(x) = \frac{\sin\alpha\pi}{\pi}\left[\frac{g(a)}{(x-a)^{1-\alpha}} + \int_a^x \frac{1}{(x-t)^{1-\alpha}}\left[\frac{d}{dt}g(t)\right]dt\right]. \tag{5}$$

Of frequent occurrence is another form, which may be called the “exterior” Abelian transform in contrast to “interior” Abelian transform (3),

$$g(x) = \int_x^{+\infty} \frac{g(t)}{(t-x)^\alpha}dt.$$

In this case, the inversion formula is

$$f(x) = -\frac{\sin\alpha\pi}{\pi}\frac{d}{dx}\int_x^{+\infty} \frac{1}{(t-x)^{1-\alpha}}g(t)dt \tag{6}$$

or, after integration by parts,

$$f(x) = -\frac{\sin\alpha\pi}{\pi}\int_{x}^{+\infty}\frac{1}{(t-x)^{1-\alpha}}\left[\frac{d}{dt}g(t)\right]dt. \qquad (7)$$

ABELIAN TRANSFORM IN TRANSMISSION TOMOGRAPHY

The integral methods of transmission computerized tomography (TCT) are based on the Radon transform [3, 4]

$$p(\xi, \theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} s(x, y)\delta(\xi - x\cos\theta - y\sin\theta)dxdy, \quad (8)$$

where $\delta(\cdot)$ is the Dirac delta.

Here, as is customary in computerized tomography, we introduce fixed, (x, y), and rotating, (ξ, ζ), coordinate systems (see Fig. 1), which are related as

$$\begin{cases}\xi = x\cos\theta + y\sin\theta\\ \zeta = -x\sin\theta + y\cos\theta.\end{cases} \qquad (9)$$

For a radially symmetric function $s(x, y) = s(\sqrt{x^2+y^2}) = s(r)$ in the polar coordinate system

$$\begin{cases}x = r\cos\varphi\\ y = r\sin\varphi,\end{cases} \qquad (10)$$

Radon transform $p(\xi, \theta)$ turns into a particular case of Abelian transform $p(\xi)$,

$$\begin{aligned} p(\xi) &= \int_{0}^{+\infty}\int_{0}^{2\pi} s(r)\delta(\xi - r\cos\varphi)rdrd\varphi \\ &= \int_{0}^{+\infty}\frac{2s(r)r}{\sqrt{r^2-\xi^2}}\eta(r - |\xi|)dr = \int_{|\xi|}^{+\infty}\frac{2s(r)r}{\sqrt{r^2-\xi^2}}dr, \end{aligned} \qquad (11)$$

where $\eta(\cdot)$ is the Heaviside unit step function.

Using general inversion formula (6) for the Abelian transform, one can write the inverse transform for (11) as

$$s(r) = -\frac{1}{\pi r}\frac{d}{dr}\int_{r}^{+\infty}\frac{\xi p(\xi)}{\sqrt{\xi^2-r^2}}d\xi. \qquad (12)$$

Integrating (12) by parts yields

$$s(r) = -\frac{1}{\pi}\int_{r}^{+\infty}\frac{1}{\sqrt{\xi^2-r^2}}\left[\frac{dp(\xi)}{d\xi}\right]d\xi. \qquad (13)$$

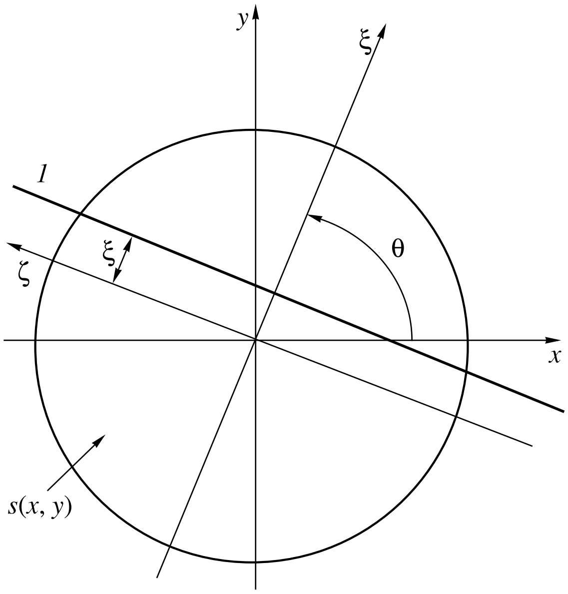


Fig. 1. Fixed, (x, y), and rotating, (ξ, ζ), coordinate systems used in the tomographic reconstruction of function $s(x, y)$. (*1*) Projection line.

Here, the regularity condition at infinity, $p(\xi) \longrightarrow 0$ as $\xi \longrightarrow +\infty$, is certainly assumed to be met.

Note that, in his classical papers [11, 12], Cormack proposed an inversion formula different from (12) and (13), namely,

$$s(r) = -\frac{1}{\pi}\frac{d}{dr}\int_{r}^{+\infty}\frac{rp(\xi)}{\xi\sqrt{\xi^2-r^2}}d\xi. \qquad (14)$$

However, integration by parts brings formula (14) to form (13), as could be expected.

Of particular interest is derivation of the inverse Abelian transform directly from the inverse Radon transform. As the initial inversion formula, let us take the one that corresponds to the filtered back projection method [3, 4],

$$\begin{aligned} &s(x, y) \\ &= \frac{1}{2\pi}\int_{0}^{2\pi}\int_{-\infty}^{+\infty} p(\xi, \theta)h(x\cos\theta + y\sin\theta - \xi)d\xi d\theta. \end{aligned} \qquad (15)$$

Here, $h(\cdot)$ is a filter function that admits an integral representation in the form of the inverse Fourier transform that exists on the class of generalized functions,

$$h(\xi) = \frac{1}{4\pi}\int_{-\infty}^{+\infty}|\chi|e^{i\chi\xi}d\xi = -\frac{1}{2\pi}\mathfrak{p}\frac{1}{\xi^2}, \qquad (16)$$

where $\mathfrak{p}(1/\xi^2)$ is a generalized function [13] that acts in accordance with the rule

$$\begin{aligned}\left(\mathfrak{p}\frac{1}{\xi^2}, f(\xi)\right) &= \int\limits_{-\infty}^{+\infty} \mathfrak{p}\frac{1}{\xi^2} f(\xi)d\xi \\ \equiv v.p.\int\limits_{-\infty}^{+\infty}\frac{f(\xi)-f(0)}{\xi^2}d\xi &= v.p.\int\limits_{-\infty}^{+\infty}\frac{f'(\xi)}{\xi}d\xi,\end{aligned} \tag{17}$$

and *v.p.* means integration in the sense of the Cauchy principal value.

Then,

$$\begin{aligned} s(x, y) &= \frac{1}{2\pi}\int\limits_{0}^{2\pi}\int\limits_{-\infty}^{+\infty} p(\xi, \theta)\left[-\frac{1}{2\pi}\mathfrak{p}\frac{1}{(x\cos\theta + y\sin\theta - \xi)^2}\right]d\xi d\theta \\ &= -\frac{1}{(2\pi)^2}\int\limits_{0}^{2\pi} v.p.\int\limits_{-\infty}^{+\infty}\left[\frac{p(\xi, \theta) - p(0, \theta)}{(x\cos\theta + y\sin\theta - \xi)^2}\right]d\xi d\theta \\ &= \frac{1}{(2\pi)^2}\int\limits_{0}^{2\pi} v.p.\int\limits_{-\infty}^{+\infty}\frac{1}{x\cos\theta + y\sin\theta - \xi}\left[\frac{\partial p(\xi, \theta)}{\partial \xi}\right]d\xi d\theta,\end{aligned} \tag{18}$$

which is the inverse Radon transform in the Radon form [5].

In the case of radial symmetry, $p(\xi, \theta) = p(\xi)$ and, passing to the polar coordinates, we have

$$\begin{aligned} s(r) &= \frac{1}{(2\pi)^2}\int\limits_{0}^{2\pi} v.p.\int\limits_{-\infty}^{+\infty}\frac{1}{r\cos\theta - \xi}\frac{dp(\xi)}{d\xi}d\xi d\theta \\ &= \frac{1}{(2\pi)^2}\int\limits_{-\infty}^{+\infty}\frac{dp(\xi)}{d\xi}\left[v.p.\int\limits_{0}^{2\pi}\frac{1}{r\cos\theta - \xi}d\theta\right]d\xi.\end{aligned} \tag{19}$$

Note that $p(\xi)$ is an even function, while $dp(\xi)/d\xi$ is an odd function. Since

$$v.p.\int\limits_{0}^{2\pi}\frac{1}{r\cos\theta - \xi}d\theta = \frac{(-2\pi)}{\sqrt{\xi^2 - r^2}}\eta(|\xi| - r)\operatorname{sgn}(\xi), \tag{20}$$

where $\operatorname{sgn}(\cdot)$ is the signum function, the inverse Radon transform yields the inverse Abelian transform in form (13),

$$\begin{aligned} s(r) &= \frac{1}{(2\pi)^2}\int\limits_{-\infty}^{+\infty}\frac{dp(\xi)}{d\xi}\left[\frac{(-2\pi)}{\sqrt{\xi^2 - r^2}}\eta(|\xi| - r)\operatorname{sgn}(\xi)\right]d\xi \\ &= -\frac{1}{\pi}\int\limits_{r}^{+\infty}\frac{1}{\sqrt{\xi^2 - r^2}}\frac{dp(\xi)}{d\xi}d\xi.\end{aligned} \tag{21}$$

EXPONENTIAL ABELIAN TRANSFORM IN EMISSION TOMOGRAPHY

The integral methods in emission computerized tomography (ECT) are based on the exponential Radon transform [3, 4]

$$\begin{aligned} p(\xi, \theta) &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} s(x, y)\delta(\xi - x\cos\theta - y\sin\theta) \\ &\quad\times e^{\mu(-x\sin\theta + y\cos\theta)}dxdy,\end{aligned} \tag{22}$$

where μ is the coordinate-independent coefficient of linear attenuation of radiation.

For instance, for the gammas from the ^{99m}Tc radionuclide with an energy of 140 keV in water, we have $\mu \approx 0.15$ cm^{-1}. For the radially symmetric function $s(x, y) = s(\sqrt{x^2 + y^2}) = s(r)$ considered above, exponential Radon transform $p(\xi, \theta)$ turns into the exponential Abelian transform $p(\xi)$,

$$\begin{aligned} p(\xi) &= \int\limits_{0}^{+\infty}\int\limits_{0}^{2\pi} s(r)\delta(\xi - r\cos\varphi)e^{\mu r\sin\varphi}rdrd\varphi \\ &= \int\limits_{0}^{+\infty}\frac{2s(r)\cosh(\mu\sqrt{r^2 - \xi^2})}{\sqrt{r^2 - \xi^2}}\eta(r - |\xi|)rdr \\ &= \int\limits_{|\xi|}^{+\infty}\frac{2s(r)\cosh(\mu\sqrt{r^2 - \xi^2})}{\sqrt{r^2 - \xi^2}}rdr.\end{aligned} \tag{23}$$

In contrast to [10], we will derive the inverse exponential Abelian transform directly from the inverse exponential Radon transform. According to the filtered back projection method [4, 8], we have

$$\begin{aligned} s(x, y) &= \frac{1}{2\pi}\int\limits_{0}^{2\pi}\int\limits_{-\infty}^{+\infty} p(\xi, \theta)h_\mu(x\cos\theta + y\sin\theta - \xi) \\ &\quad\times e^{-\mu(-x\sin\theta + y\cos\theta)}d\xi d\theta,\end{aligned} \tag{24}$$

where $h_\mu(\cdot)$ is a filter function that also has an integral representation in the form of the inverse Fourier transform that exists on the class of generalized functions:

$$\begin{aligned} h_\mu(\xi) &= \frac{1}{4\pi}\int\limits_{-\infty}^{+\infty}|\chi|[1 - \eta(\mu - |\chi|)]e^{i\chi\xi}d\chi \\ &= -\frac{1}{2\pi}\mathfrak{p}\frac{1}{\xi^2} - \frac{\mu^2}{2\pi}\left[\operatorname{sinc}(\mu\xi) - \frac{1}{2}\operatorname{sinc}^2\left(\frac{\mu\xi}{2}\right)\right].\end{aligned} \tag{25}$$

Here, $\operatorname{sinc}(x) \equiv \sin x/x$.

Then, the inverse exponential Radon transform can be written in the Radon form as

$$s(x, y) = \frac{1}{2\pi}\int_0^{2\pi}\int_{-\infty}^{+\infty} p(\xi, \theta)\left\{-\frac{1}{2\pi}\mathfrak{p}\frac{1}{(x\sin\theta + y\sin\theta - \xi)^2}\right.$$

$$-\frac{\mu^2}{2\pi}\operatorname{sinc}[\mu(x\cos\theta + y\sin\theta - \xi)]$$

$$\left.+\frac{\mu^2}{4\pi}\operatorname{sinc}^2\left[\frac{\mu(x\cos\theta + y\sin\theta - \xi)}{2}\right]\right\} \tag{26}$$

$$\times e^{-\mu(-x\sin\theta + y\cos\theta)}d\xi d\theta$$

$$= \frac{1}{(2\pi)^2}\int_0^{2\pi} v.p.\int_{-\infty}^{+\infty}\frac{\cos[\mu(x\cos\theta + y\sin\theta - \xi)]}{x\cos\theta + y\sin\theta - \xi}$$

$$\times e^{-\mu(-x\sin\theta + y\cos\theta)}\left[\frac{\partial p(\xi, \theta)}{\partial\xi}\right]d\xi d\theta.$$

In the case of radial symmetry, $p(\xi, \theta) = p(\xi)$ and, passing to polar coordinates (10), we find

$$s(r) = \frac{1}{(2\pi)^2}$$

$$\times\int_0^{2\pi} v.p.\int_{-\infty}^{+\infty}\frac{\cos[\mu(r\cos\theta - \xi)]}{r\cos\theta - \xi}e^{\mu r\sin\theta}\frac{dp(\xi)}{d\xi}d\xi d\theta \tag{27}$$

$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\left[v.p.\int_0^{2\pi}\frac{\cos[\mu(r\cos\theta - \xi)]}{r\cos\theta - \xi}e^{\mu r\sin\theta}d\theta\right]\frac{dp(\xi)}{d\xi}d\xi.$$

It can be demonstrated that

$$v.p.\int_0^{2\pi}\frac{\cos[\mu(r\cos\theta - \xi)]}{r\cos\theta - \xi}e^{\mu r\sin\theta}d\theta$$

$$= \frac{(-2\pi)\cos(\mu\sqrt{\xi^2 - r^2})}{\sqrt{\xi^2 - r^2}}\eta(|\xi| - r)\operatorname{sgn}(\xi). \tag{28}$$

Finally, we arrive at the inverse exponential Abelian transform in the form

$$s(r) = -\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\frac{\cos(\mu\sqrt{\xi^2 - r^2})}{\sqrt{\xi^2 - r^2}}\eta(|\xi| - r)\operatorname{sgn}(\xi)\right]$$

$$\times\frac{dp(\xi)}{d\xi}d\xi = -\frac{1}{\pi}\int_r^{+\infty}\frac{\cos(\mu\sqrt{\xi^2 - r^2})}{\sqrt{\xi^2 - r^2}}\frac{dp(\xi)}{d\xi}d\xi. \tag{29}$$

NUMERICAL IMPLEMENTATION OF THE INVERSE EXPONENTIAL ABELIAN TRANSFORM

When numerically implementing the inverse exponential Abelian transform, one faces a number of problems that introduce distortions into the results obtained. Among such distortions are, first of all, those due to the ill-posedness inherent in the problem of solving integral equations of the second kind. In practice, this makes the results unstable when the right-hand side has a statistic error. In this paper, no consideration is given to such distortions, since, on the one hand, they have been extensively discussed in the literature (see, e.g., [14]) and, on the other hand, this topic calls for separate investigation.

Distortions of another type are due to the structure of the object itself; e.g., they may arise because of discontinuities at the boundaries of or inside the object (Fig. 2). Such distortions can be effectively corrected by considering each of them as an additive component proportional to the amount of a discontinuity. Such a correlation works well if the statistical noise is low; if it is high, general methods of regularization of ill-posed problems in combination with distortion correction techniques should be applied.

Next, distortions may be associated with intrinsic peculiarities of a particular problem. In the case of the exponential Abelian transform, the loss of information from deep-seated regions of the object is an example. For the conventional Abelian transform, when $\mu = 0$, this problem does not arise. The Fourier image of filter

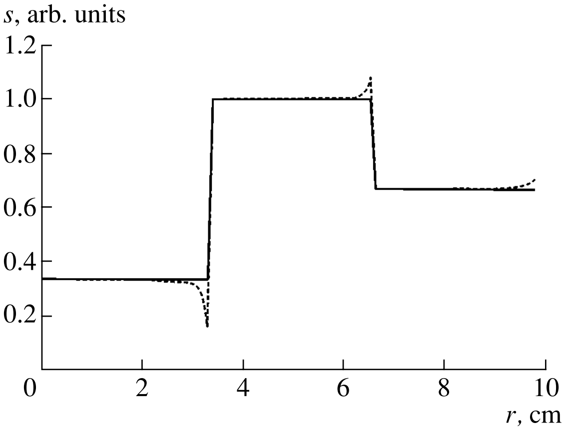


Fig. 2. Correction of the distortions associated with discontinuities in the radiation source distribution function. The solid line shows the original function; the dashed line, the function reconstructed without correction. The function reconstructed with correction coincides with the original function.

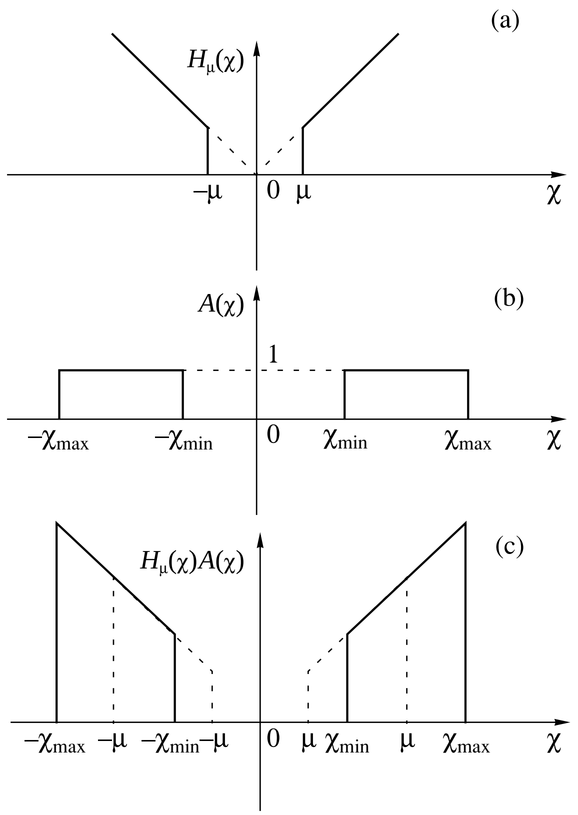


Fig. 3. (a) Fourier image of the filter function, (b) apodizing function, and (c) Fourier image of the filter function with regard to the apodizing function.

function (25),

$$H_\mu(\chi) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} h_\mu(\xi)e^{-i\chi\xi}d\xi = \frac{1}{2\sqrt{2\pi}}|\chi|\eta(|\chi|-\mu), \qquad (30)$$

is shown in Fig. 3a. In calculations, function $H_\mu(\chi)$ should be multiplied by some apodizing function $A(\chi)$ that regularizes the ill-posed problem of finding the inverse EAT. Even without the apodizing function, the maximal size of domain R being discretized and discretization interval Δr specify the lowest, χ_{min}, and highest, χ_{max}, spatial frequencies in the Fourier space. In essence, this corresponds to introduction of the apodizing function presented in Fig. 3b. Then, with regard to the apodizing function, the Fourier image of the filter function depends on a relationship between μ and χ_{min} (see Fig. 3c). If $\mu < \chi_{min}$, the lowest of low spatial frequencies is equal to χ_{min}; for $\mu > \chi_{min}$, it equals μ. In the latter case, information is partially lost and the object reconstructed is distorted. Physically, this means that the radiation from very deep-seated areas of the object is not detected at all. Since the lowest spatial frequency corresponds to the maximal size of the object, attention should be paid to a proper relationship between the object size and the radiation attenuation coefficient.

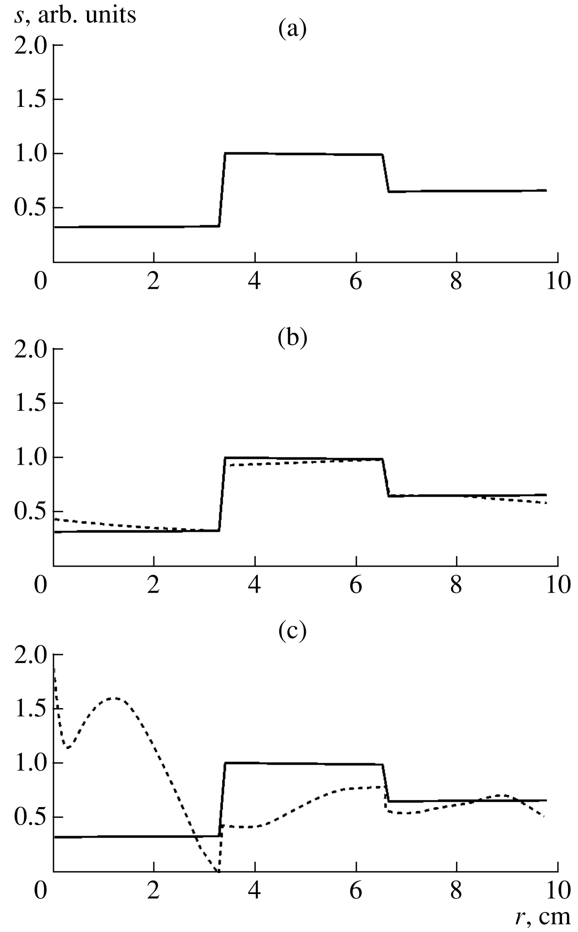


Fig. 4. Effect of the radiation attenuation coefficient on the quality of reconstruction of the radiation source distribution function. The solid line, the original function; the dashed line, the function reconstructed. μ = (a) $0.5\chi_{min} = 0.005\chi_{max}$, (b) $5\chi_{min} = 0.05\chi_{max}$, and (c) $10\chi_{min} = 0.1\chi_{max}$.

The aforesaid is illustrated by Fig. 4. As long as $\mu < \chi_{min}$, the original function is restored almost exactly (Fig. 4a). When $\mu > \chi_{min}$, the function reconstructed contains distortions (Fig. 4b), which build up as the truncated part of the spatial frequency band expands with μ. Even if this part is about 10%, the distortions become unallowably large (Fig. 4c).

CONCLUSIONS

The basic properties of the exponential Abelian transform as a generalization of the classical Abelian transform are considered. A new approach to deriving the inverse exponential Abelian transform directly from the inverse exponential Radon transform is suggested. The problems of numerical implementation of the inverse exponential Abelian transform are discussed. The principal distinction between the exponential Abelian transform and the conventional Abelian transform is that the radiation (information) from very deep-seated regions of the object is partially lost in the former case. The results obtained may be useful in reconstructing the spatial distribution of axisymmetric or spherically symmetric radiation sources, e.g., in diagnostics of fuel elements for nuclear reactors, as well as of other objects.

ACKNOWLEDGMENTS

The work was partially supported by the Russian Foundation for Basic Research, grant no. 02-02-96000.

REFERENCES

1. G. T. Herman, *Image Reconstruction from Projection: The Fundamentals of Computerized Tomography* (Academic, New York, 1980; Mir, Moscow, 1983).
2. G. G. Levin and G. N. Vishnyakov, *Optical Tomography* (Radio i Svyaz', Moscow, 1989) [in Russian].
3. F. Natterer, *The Mathematics of Computerized Tomography* (Wiley, Chichester, 1986; Mir, Moscow, 1990).
4. G. A. Fedorov and S. A. Tereshchenko, *Computerized Emission Tomography* (Énergoatomizdat, Moscow, 1990) [in Russian].
5. S. Helgason, *The Radon Transform* (Birkhuser, Boston, 1980; Mir, Moscow, 1983), pp. 134–148.
6. S. Bellini, M. Piacentini, C. Cafforio, *et al.*, IEEE Trans. Acoust., Speech, Signal Process. **27**, 213 (1979).
7. O. Tretiak and C. Metz, SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **39**, 341 (1980).
8. G. T. Gullberg and T. F. Budinger, IEEE Trans. Biomed. Eng. **28**, 142 (1981).
9. *Abelian Inversion and Its Generalizations*, Ed. by N. G. Preobrazhenskiĭ (ITPM, Novosibirsk, 1978) [in Russian].
10. A. V. Clough and H. H. Barrett, J. Opt. Soc. Am. **73**, 1590 (1983).
11. A. M. Cormack, J. Appl. Phys. **34**, 2722 (1963).
12. A. M. Cormack, J. Appl. Phys. **35**, 2908 (1964).
13. V. S. Vladimirov, *Equations of Mathematical Physics* (Nauka, Moscow, 1981; Dekker, New York, 1971).
14. A. N. Tikhonov and V. Ya. Arsenin, *Solutions of Ill-Posed Problems* (Nauka, Moscow, 1986; Halsted, New York, 1977).

Translated by A. Pankrat'ev

Technical Physics, Vol. 50, No. 2, 2005, pp. 158–166. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 19–27.

GASES
AND LIQUIDS

Nonlinear Analysis of the Time Evolution of an Unstable Flat Charged Fluid Surface

A. I. Grigor'ev, S. O. Shiryaeva, D. F. Belonozhko, and A. V. Klimov
Yaroslavl State University, Sovetskaya ul. 14, Yaroslavl, 150000 Russia
e-mail: grig@uniyar.ac.ru

Received October 27, 2003; in final form, January 13, 2004

Abstract—Nonlinear second- and fourth-order corrections to the critical Tonks–Frenkel parameter (which characterizes the stability of the uniformly charged flat surface of an ideal conducting incompressible fluid) are found by asymptotic calculations of the fifth order of smallness in ratio of the wave amplitude to the capillary constant of the fluid. A nonlinear integral equation for the time evolution of the unstable wave amplitude is derived and solved. It turns out that the linear stage of instability development takes a major part of the total time, while the nonlinear stage is very short. It is shown that the characteristic time of instability development on the fluid surface is a rapidly decreasing function of the initial amplitude of a virtual wave and the overcritical surface charge (i.e., the excess of the charge over the critical value). *© 2005 Pleiades Publishing, Inc.*

(1) Investigation into the physical mechanisms of instability development on the charged flat fluid surface is of both scientific and applied interest (see, for example, [1–9] and Refs cited therein). Most of the theoretical studies has been conducted in an approximation linear in amplitude of deformation of the flat fluid surface. Works reflecting the nonlinear essence of this phenomenon have appeared only recently [10–15]. Yet, many issues concerning its nature still remain unclear. In particular, mechanisms behind the formation of Taylor cones have not been covered to date. By Taylor cones, researchers mean protrusions arising on the charged fluid surface at the nonlinear stage of instability development. These protrusions emit finely dispersed highly charged droplets from their tops, thereby removing an extra charge [1–8]. A qualitative model of protrusion formation was proposed by Tonks as early as in 1936 [1]. Numerical analysis of the protrusions was discussed in [10, 16]. However, nobody has tried to evaluate the characteristic time of Taylor cone formation from the very beginning of instability development on the charged fluid surface. The instability growth rate as a parameter characterizing the growth time of a protrusion is inappropriate. The fact is that it adequately describes the growth of the protrusions only at the linear (initial) stage, i.e., as long as the height (amplitude) of a protrusion remains much less than its transverse linear size, whereas a Taylor cone is an essentially nonlinear object. In this work, we calculate the characteristic time of Taylor cone formation following the scheme used earlier in the analysis of the nonlinear stages of the evolution of a highly charged drop [17, 18] and of an uncharged drop subjected to a high uniform external electrostatic field [19].

(2) Let an ideal perfectly conducting incompressible fluid with a density ρ occupy the space $z \le 0$ in the Cartesian coordinate system. The fluid experiences the action of a gravitational field ($\mathbf{g} \parallel -\mathbf{n}_z$, where $\mathbf{n}_z$ is the unit vector in the direction of the z axis). The equilibrium (undisturbed by wave motion) flat surface of the fluid (which coincides with plane XOY) borders a vacuum, experiences the action of surface tension forces with a surface tension coefficient γ, and bears a uniformly distributed electric charge with a density σ. The wave motion of fluid molecules generates small-amplitude capillary waves on the free fluid surface, so that its shape becomes a function of coordinate and time: $z = \xi(x, t)$. The mathematical statement of the problem of nonlinear capillary–gravitational waves on the free uniformly charged fluid surface has the form

$$z > \xi:\ \Delta\Phi = 0; \quad z \le \xi:\ \Delta\varphi = 0;$$

$$z = \xi:\ p = p_0 - \rho g z - \rho\frac{\partial\varphi}{\partial t} - \frac{\rho}{2}\left(\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2\right);$$

$$\frac{\partial\xi}{\partial t} + \frac{\partial\xi}{\partial x}\frac{\partial\varphi}{\partial x} = \frac{\partial\varphi}{\partial z};$$

$$p + \frac{(\nabla\Phi)^2}{8\pi} = -\gamma\frac{\partial^2\xi}{\partial x^2}\left(1 + \left(\frac{\partial\xi}{\partial x}\right)^2\right)^{-3/2}; \quad \Phi = 0;$$

$$z \longrightarrow \infty:\ \nabla\Phi \longrightarrow -E_0\mathbf{e}_z;$$

$$z \longrightarrow -\infty:\ \nabla\phi \longrightarrow 0.$$

Here, $\varphi(\mathbf{r}, t)$ is the potential of the fluid velocity field, $\Phi(\mathbf{r}, t)$ is the potential of the electrostatic field above the fluid, $p(\mathbf{r}, t)$ is the pressure field in the fluid, and Δ is the Laplacian.

It is also necessary to formulate the initial conditions of the problem. The solution of nonlinear problems with arbitrarily postulated initial conditions is a challenge. Therefore, it is reasonable to specify the initial conditions in such a way that the result of the solution has a form that is as simple as possible. To this end, we suppose that a capillary–gravitational wave on the fluid surface is a travelling sinusoidal wave in a first approximation in small wave amplitude; that is, its profile has the form

$$\xi = a\cos(kx - \omega t) + O(a^2),$$

where a and ω are the amplitude and frequency of the wave, respectively, and k is the wavenumber.

We assume that the wave amplitude is much smaller than the capillary constant of the fluid, $\alpha = \sqrt{\gamma/\rho g}$, so that ratio a/α is natural to take as a small parameter in the following asymptotic analysis. It is also assumed that all second- and higher order additions to the wave profile, which are proportional to $\cos(kx - \omega g)$ and may appear due to the nonlinearity of the problem, are equal to zero.

In a zeroth approximation in small parameter a/α, the free fluid surface is unperturbed and described by the equation $z = 0$, the fluid is quiescent, and the electric field is uniform throughout the space:

$$\xi_0 \equiv 0; \quad \nabla\varphi_0 \equiv 0; \quad \nabla\Phi_0 \equiv E_0\mathbf{e}_z; \quad p_0 = -\frac{E_0^2}{8\pi}.$$

Substituting these expressions into the initial equations gives

$$\Phi_0 \equiv -E_0 z.$$

The perturbation ξ of the free surface, velocity field potential φ, and electric potential Φ are unknown functions in the problem. We shall seek them in the form of expansions in small parameter,

$$\xi = \left(\frac{a}{\alpha}\right)\xi_1 + \left(\frac{a}{\alpha}\right)^2\xi_2 + \left(\frac{a}{\alpha}\right)^3\xi_3 + \left(\frac{a}{\alpha}\right)^4\xi_4 + O\left(\frac{a}{\alpha}\right)^5;$$

$$\Phi = -E_0 z + \left(\frac{a}{\alpha}\right)\Phi_1 + \left(\frac{a}{\alpha}\right)^2\Phi_2 + \left(\frac{a}{\alpha}\right)^3\Phi_3 + \left(\frac{a}{\alpha}\right)^4\Phi_4 + O\left(\frac{a}{\alpha}\right)^5;$$

$$\varphi = \left(\frac{a}{\alpha}\right)\varphi_1 + \left(\frac{a}{\alpha}\right)^2\varphi_2 + \left(\frac{a}{\alpha}\right)^3\varphi_3 + \left(\frac{a}{\alpha}\right)^4\varphi_4 + O\left(\frac{a}{\alpha}\right)^5;$$

$$\xi_n \sim O(1), \quad \Phi_n \sim O(1), \quad \varphi_n \sim O(1).$$

(3) We will solve the problem by the method of many scales [14, 20–22] in the fifth order of smallness in wave amplitude using dimensionless variables such that $g = \gamma = \rho = 1$ (consequently, the capillary constant of the fluid is also equal to unity, $\alpha = 1$). Then, the profile of a capillary–gravitational wave will take the form

$$\begin{aligned}\xi = {} & a\cos[kx - (\omega + a^2\delta_2 + a^4\delta_4)t] \\ & + a^2X_{22}\cos[2kx - 2(\omega + a^2\delta_2)t] \\ & + a^3X_{33}\cos[3kx - 3(\omega + a^2\delta_2)t] \\ & + a^4(X_{42}\cos[2(kx - \omega t)] + X_{44}\cos[4(kx - \omega t)]) \\ & + a^5X_5\cos[5(kx - \omega t)];\end{aligned} \tag{1}$$

(all variables in (1) are designated as before, except for the small parameter, which is denoted by a). Here,

$$\omega^2 = k(1 + k^2 - kW); \quad W = 4\pi\sigma^2;$$

$$\delta_2 = \frac{k^3(16k^2W^2 - 16k^3W + 2k^4 - 16kW + k^2 + 8)}{16\omega(1 - 2k^2)};$$

$$\begin{aligned}\delta_4 = {} & \frac{k^5}{512\omega(1 + k^2 - kW)(1 - 2k^2)^3(1 - 3k^2)} \\ & \times(256 - 2386kW + 788k^2 + 8704k^2W^2 - 5188k^3W \\ & - 15616k^3W^3 + 229k^4 + 11040k^4W^2 + 13056k^4W^4 \\ & + 942k^5W - 2432k^5W^3 - 4096k^5W^5 - 377k^6 \\ & - 18528k^6W^2 - 9984k^6W^4 + 11948k^7W + 31232k^7W^3 \\ & + 6144k^7W^5 - 2218k^8 - 22400k^8W^2 - 13824k^8W^4 \\ & + 6056k^9W + 7680k^9W^3 - 1148k^{10} + 1920k^{10}W^2 \\ & - 1968k^{11}W + 24k^{12});\end{aligned}$$

$$X_{22} = \frac{k(1 + k^2 - 2kW)}{2(1 - 2k^2)};$$

$$X_{33} = \frac{k^2(6 - 32kW + 21k^2 + 32k^2W^2 - 32k^3W + 6k^4)}{16(1 - 2k^2)(1 - 3k^2)};$$

$$\begin{aligned}X_{42} = {} & \frac{k^3}{48(1 - 2k^2)^3(1 - 3k^2)}(34 - 164kW - 33k^2 \\ & + 336k^2W^2 - 52k^3W - 192k^3W^3 + 171k^4 \\ & - 240k^4W^2 + 124k^5W + 192k^5W^3 \\ & - 134k^6 - 24k^7W - 48k^8);\end{aligned}$$

$$X_{44} = \frac{k^3(1+k^2-2kW)}{48(1-2k^2)^2(1-3k^2)(1-4k^2)}(16-128kW$$
$$+77k^2+128k^2W^2+160k^3W-248k^4-288k^4W^2$$
$$+288k^5W-12k^6),$$

and W is the dimensionless Tonks–Frenkel parameter characterizing the stability of the free fluid surface against the surface charge.

The expression for X_5 is omitted, because it is awkward and has an insignificant influence on the profile of the resulting wave. In the limit $W \longrightarrow 0$ (i.e., $\sigma \longrightarrow 0$), expression (1) for the shape of a nonlinear wave on the free charged surface of an ideal fluid coincides with the known expression [21, 22] for the shape of nonlinear capillary–gravitational waves on the uncharged surface of an ideal fluid up to the third order of smallness.

From (1), it is seen that amplitude factor X_{22} multiplying the second-order correction builds up resonantly at $k = k_2 = 1/(2^{1/2})$ and factor X_{33} multiplying the third-order correction, at $k = k_2$ and $k = k_3 = 1/(3^{1/2})$. The amplitude factor X_{44} of the fourth-order correction has three resonances: at $k = k_2$, $k = k_3$, and $k = k_4 = 1/(4^{1/2})$. It was shown [20] that, in an approximation quadratic in wave amplitude (when there is only one resonant wave number $k = k_2$), the resonant interaction causes energy transfer from longer waves with wavenumbers $k = k_2$ to shorter ones with $k = 2k_2$. It is also seen from expression (1) that energy transfer takes place not only under the third-order resonance conditions mentioned above but also at $k = k_3$: the energy is transferred from longer waves with wavenumbers $k = k_3$ to shorter waves with $k = 3k_3$. In the fourth order of smallness, one more resonance is observed, namely, at $k = k_4$. Here, the energy is transferred from the waves with $k = k_4$ to those with $k = 4k_4$.

It is also seen that the frequency correction that is proportional to δ_2 involves the amplitude squared and the frequency correction proportional to δ_4, the amplitude in the fourth power. These corrections influence the frequencies in the third and fifth orders of smallness, respectively (which can be shown by expanding $a\cos[(\omega + a^2\delta_2 + a^4\delta_4)t]$ in powers of $a^2\delta_2$ and $a^4\delta_4$).

The nonlinear corrections to the frequency, as well as amplitude factors X_{22}, X_{33}, X_{42}, and X_{44}, have resonant form (i.e., their denominators have factors vanishing at certain wavenumbers). This means that the applicability of expression (1) near wavenumbers $k = k_2$, $k = k_3$, and $k = k_4$ is limited, since the amplitude factors multiplying frequency corrections δ_2 and δ_4 and amplitude coefficients X_{22}, X_{33}, X_{42}, and X_{44} must be on the order of $O(1)$.

The fundamental frequency is equal to $(\omega + a^2\delta_2 + a^4\delta_4)$. The critical conditions under which the free fluid surface is unstable against self-charge are the following: (i) vanishing of the square of the frequency Z, $Z^2 = (\omega + a^2\delta_2 + a^4\delta_4)^2$, of the virtual wave and (ii) vanishing of the derivative of frequency Z with respect to wave number. The first condition yields critical Tonks–Frenkel parameter W_*; the second, the wavenumber k_* of the most unstable wave [2, 23]. In the framework of the linear model, the critical values of W_* and k_* are related as [20]

$$W_* = k_* + k_*^{-1}, \quad k_* = 1. \tag{2}$$

In the nonlinear problem being solved, we will seek for nonlinear corrections to Tonks–Frenkel parameter W that is critical in terms of instability development under the assumption that condition (2) is met and the wave with $k = 1$ becomes unstable. To do this, we represent W in the form of the expansion

$$W_* \approx 2 - w_2a^2 + w_4a^4. \tag{3}$$

Substituting this expansion into the equation $Z^2 = 0$, one can readily obtain by the method of successive approximations that $w_2 = 11/10$ and $w_4 = 51/160$. Thus, it turns out that the critical value of parameter W at which the flat charged surface of an ideal fluid becomes unstable depends on the amplitude of the virtual wave, which loses stability, giving rise to Tonks–Frenkel instability. In full-scale experiments, the free fluid surface exhibits an infinite spectrum of capillary waves with an amplitude on the order of $\sqrt{\kappa T/\gamma}$ (κ is the Boltzmann constant, T is the absolute temperature of the fluid), which are generated by the thermal motion of fluid molecules. It is these waves that play the role of virtual waves [2]. It is easy to see that, for most real fluids, the amplitude of such waves at reasonable temperatures (i.e., at which the fluid exists) amounts to half an angstrom or, in our dimensionless variables, ~10^{-8}. In terms of instability development on the charged flat fluid surface, such a small wave amplitude means that the critical instability conditions follow from the linear theory: $k_* = 1$ and $W_* = 2$. As for the nonlinear corrections (see (3)), their influence shows up as an increase in the amplitude growth rate, which is amplitude-dependent. This phenomenon is discussed in detail below.

(4) Assume that, at the zero time ($t = 0$), parameter W at the uniformly charged free fluid surface is equal to the critical value predicted by the linear theory, $W = 2$. Then, according to the linear theory, the squared frequency of the wave with $k = 1$ and an infinitesimal amplitude vanishes [2],

$$\omega^2 = k(1 + k^2 - kW). \tag{4}$$

This circumstance as such does not yet mean that the wave becomes unstable, because, according to the linear theory, the instability growth rate of such a wave is equal to zero. The fact is that, in the linear theory, the dispersion relation does not contain the wave amplitude and all the components of the dispersion relation are of the zeroth order of smallness. The nonlinear analysis performed above shows, however, that the critical value of parameter W depends on the wave amplitude and decreases with increasing amplitude according to (3). This means that the electric field near the fluid surface (which is characterized by parameter W) exceeds the critical value and, hence, that the instability growth rate for a wave with an amplitude as small as desired is other than zero. The excess of parameter W over its critical value is defined by the difference between $W_* \approx 2 - w_2a^2 + w_4a^4$ and $W = 2$. In other words, the squared frequency of the wave is given by

$$\omega^2 \approx -w_2a^2 + w_4a^4, \tag{5}$$

that is, the wave becomes unstable with an instability growth rate χ, where $\chi = a\sqrt{(w_2 - w_4a^2)}$ according to (5). If one takes into consideration that the amplitude of the unstable wave increases with time, this relationship gives the growth rate as a function of amplitude $a(t)$, which increases with time:

$$\chi = a(t)\sqrt{(w_2 - w_4a(t)^2)}. \tag{6}$$

It should be emphasized that relationship (3) (which follows from the nonlinear analysis) is used for deriving an analytical expression for the linear parameter characterizing the time evolution of the unstable wave, i.e., the growth rate.

Since the amplitude of the wave is a function of time, the instability growth rate also depends on time (see (6)), $\chi = \chi(t)$. This means that the amplitude increases with time as

$$\begin{aligned} a(t) &= a_0\exp(\chi t) \\ &= a_0\exp[a(t)\sqrt{(w_2 - w_4a(t)^2)}t], \end{aligned} \tag{7}$$

that is, considerably faster than by the normal exponential law, and, consequently, will rapidly go beyond the domain of applicability of the expansions in small parameter that were used in the derivation of (7).

To find an amplitude variation law applicable for any amplitude values, consider a sequence of amplitude values a taken in time steps Δt_i ($i = 1, 2, 3, 4, \ldots$) such that expression (7) applies within each of them. It should be noted that (7) is valid only if amplitude growth rate da/dt is linearly related to amplitude $a(t)$. The width of step Δt_i will be found from the condition that an increase in the growth rate increment within this step, $\chi_i - \chi_{i-1}$, is much smaller than the increment at the end of the preceding step; that is, $\Delta\chi \equiv \chi_i - \chi_{i-1} \ll \chi_{i-1}$. Then, the value of χ_i may be considered invariable within step Δt_i and the increment Δa_i of the amplitude may be calculated with relationship (7).

Let, at $W = 2$, the wave with $k = 1$ become unstable and its amplitude start to increase with growth rate $\chi_1 = a_0\sqrt{(w_2 - w_4a_0^2)}$. For a time Δt_1, the amplitude $a_0 = \sqrt{\kappa T}$ of the initial virtual wave increases to $a_1 = a_0\exp(\chi_1\Delta t_1)$. According to (6), the growth rate will then increase to $\chi_2 = a_0\sqrt{(w_2 - w_4a_1^2)}$. Within the next time step Δt_2, the amplitude will grow with such a rate to $a_2 = a_1\exp(\chi_2\Delta t_2)$. At the end of step Δt_2, the growth rate will be $\chi_3 = a_2\sqrt{(w_2 - w_4a_2^2)}$. Within the following time step Δt_3, the amplitude will grow by the law $a_3 = a_2\exp(\chi_3\Delta t_3)$.

Thus, for an ith time step Δt_i, we obtain

$$a_i = a_{i-1}\exp(\chi_i\Delta t_i). \tag{8}$$

Now, we substitute into (8) the expression for a_{i-1}, which is expressed via a_{i-2}. The latter, in its turn, is expressed via a_{i-3} and so on down to a_0. Eventually, instead of (8) we will have

$$a_i = a_0\exp[\Sigma_{m=1}^{i}(\chi_m\Delta t_m)].$$

Passing to the limit $\Delta t_m \longrightarrow 0$ ($m \longrightarrow \infty$) in this expression, we arrive at

$$a(t) = a_0\exp\left[\int_0^t \chi(t)dt\right].$$

Let us substitute into this expression relationship (6), which relates the growth rate to the amplitude:

$$a(t) = a_0\exp\left[\int_0^t a(t)\sqrt{(w_2 - w_4a(t)^2)}dt\right]. \tag{9}$$

Thus, we have derived the nonlinear integral equation for time-varying wave amplitude $a(t)$. To find a solution to Eq. (9), we take the logarithm of (9),

$$\ln(a(t)/a_0) = \int_0^t a(t)\sqrt{(w_2 - w_4a(t)^2)}dt$$

and differentiate the resulting expression with respect to t,

$$\frac{da(t)}{a(t)dt} = a(t)\sqrt{(w_2 - w_4a(t)^2)}.$$

Separating the variables, we get

$$\frac{da}{a^2\sqrt{(w_2 - w_4a^2)}} = dt.$$

Integration yields

$$\frac{\sqrt{(w_2 - w_4 a_0^2)}}{w_2 a_0} - \frac{\sqrt{(w_2 - w_4 a(t)^2)}}{w_2 a(t)} = t$$

or

$$a(t) = \left(\sqrt{w_2\left(\frac{\sqrt{(w_2 - w_4 a_0^2)}}{w_2 a_0} - t\right)^2 + \frac{w_4}{w_2}}\right)^{-1}. \qquad (10)$$

From expression (10), it follows that the characteristic time τ of instability development (the characteristic time of nonlinear increase of the amplitude, which can be determined as the time at which expression (10) reaches a maximum) has the form

$$\tau = \frac{\sqrt{(w_2 - w_4 a_0^2)}}{w_2 a_0}. \qquad (11)$$

Considering the instability development of the virtual wave (beginning from its thermal amplitude at $W = 2$), we can argue that the characteristic time of amplitude growth is very long. Specifically, in terms of our dimensionless variables, it reaches $\sim 10^8$; in the dimensional variables, for example, for water bordering a vacuum, this time is as long as 2.5×10^4 s, i.e., about 7 h (the characteristic scale of dimensionless time is $\sqrt{\rho g^3/\gamma}$; that is, $\tau \equiv t\sqrt{\rho g^3/\gamma}$). For other fluids or their mixtures, this time may somewhat change depending on the interfacial tension coefficient, densities of the media, and temperature. However, the order of magnitude of the characteristic time remains the same, since just the physical parameters listed above influence the thermal amplitude of virtual waves and the scales on which the quantities are made dimensionless. This is in qualitative agreement with the observations of Taylor and McIwan [8], who noted that the linear stage of instability development is many times longer than the nonlinear stage, during which the amplitude of emitting protrusions (Taylor cones) on the charged fluid surface rapidly increases. Nevertheless, the authors of experimental works [3, 24] did not report a long delay of instability development, although this fact is hard to escape the researchers' attention. This suggests that, in those experiments, instability developed on the virtual wave, whose amplitude is considerably larger than that of the wave due to the thermal motion of fluid molecules.

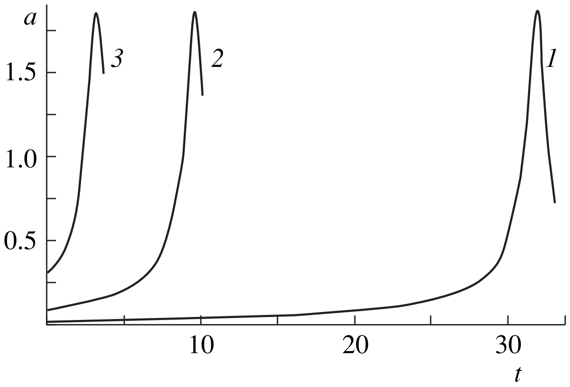


Fig. 1. Dimensionless amplitudes of nonlinearly growing waves with the dimensionless wavenumber $k = 1$ as functions of dimensionless time (expression (10)) for $W = 2$ and initial amplitudes a_0 = (*1*) 0.03, (*2*) 0.10, and (*3*) 0.30.

Figure 1 demonstrates the time dependencies of the amplitudes of nonlinearly increasing waves that were calculated at various initial amplitudes far exceeding the thermal one. Along with the delay of instability development mentioned above, one can also see that the amplitude of a nonlinearly increasing wave passes through a maximum and then decreases. The maximum amplitude of the nonlinear wave is independent of its initial amplitude but depends on a relationship between the coefficients multiplying the second- and fourth-order nonlinear corrections. This is because these corrections to the critical Tonks–Frenkel parameter have opposite signs (see (3)). If the second-order correction alone is taken into consideration, the limitation on the amplitude (see Fig. 1) is removed and the amplitude grows to infinity. This circumstance is most likely to be associated merely with a limited applicability of the relationships derived; that is, they rapidly fall outside the domain where expansion (3) is uniformly applicable. As the amplitude grows, the fourth- and second-order corrections become equal to each other, causing the maxima in the curves and then (at long times) the former correction is responsible for amplitude damping.

(5) One more factor causing the unstable wave to rapidly build up is the overcritical electric field strength (surface charge density σ) applied to the unperturbed fluid surface. In experiments, the critical conditions for instability development are usually stated by relationship (2) for well-conducting fluids. In real well-conducting fluids, the time of Maxwell relaxation of an electric charge must be much shorter than the characteristic time of instability development in order to prevent the electric potential of the fluid surface from being equalized faster than the surface deformation is complete. In most of the experiments known (see, e.g., [8, 24]), this requirement was met but the critical instability conditions differed from (2) and depended on the experimental conditions (in [24], the critical conditions depend on the characteristic time of rise of the electric field strength near the fluid surface). This allows us to suppose that the electric field in the experiments [8, 24] somewhat exceeded the critical value $W_* = 2$; consequently, the instability growth rate of the virtual wave was governed by two factors: the excess of parameter W over its critical value and nonlinear corrections (3). Therefore, we will repeat the line of reasoning similar to that in Sect. 4 of this paper, taking into account that

Tonks–Frenkel parameter W exceeds its critical value W_* by ΔW and, thus, contributes to the growth rate. In this case, instead of (6), the instability growth rate of the charged surface of an ideal incompressible fluid will take the form

$$\chi = \sqrt{\Delta W + a(t)^2(w_2 - w_4 a(t)^2)}.$$

Then, instead of (9), the amplitude of the unstable wave as a function of time is given by the nonlinear integral equation

$$a(t) = a_0 \exp\left[\int_0^t \sqrt{\Delta W + a(t)^2(w_2 - w_4 a(t)^2)}\,dt\right],$$

which has the solution

$$a(t) = \sqrt{\frac{4\Delta W F(t)}{1 + (w_2^2 + 4\Delta W w_4)F(t)^2 - 2w_2 F(t)}}, \quad (12)$$

where

$$F(t) = \frac{a_0^2 \exp(\sqrt{\Delta W}t)}{2\Delta W + w_2 a_0^2 + 2\sqrt{\Delta W(\Delta W + w_2 a_0^2 - w_4 a_0^2)}}.$$

Figure 2 plots the amplitudes of the nonlinearly increasing waves against time that were calculated by (12) at various $\Delta W = W - W_*$ and the initial thermal amplitude $a_0 = 10^{-8}$ of the virtual waves. It is easy to see that the curves in Fig. 2 are qualitatively similar to those presented in Fig. 1. The discrepancy is only quantitative: as was mentioned above, the characteristic time of instability development at $\Delta W = 0$ and $a_0 = 10^{-8}$ is very large, $t \sim 10^8$, while, at high $W \neq 0$ and $a_0 = 10^{-8}$, the dimensionless characteristic time of instability development may be very short, $t = 10$–100. The latter value is in good agreement with the experimental data in [8, 24]. However, the accuracy of measurement of the electric field in [8, 24] is doubted, since there is reason to think that parameter W in the experiments [8, 24] exceeded the critical value by a certain value ΔW. It is seen from Fig. 2 that variation of excess ΔW from several percent to 100% of W_* influences the process phenomenology (the instability development time) only slightly. In other words, only when the characteristic time of instability development is ≈2.5×10^4 s (i.e., when the instability at $W = W_* = 2$ is due to virtual waves of thermal nature) can measurements of the critical conditions for instability development on the charge surface of a conducting fluid be considered correct. Only in this case does it make sense to believe that the true field strength is measured and that the wave with $k = \alpha^{-1}$ becomes unstable, as follows from the theory [2].

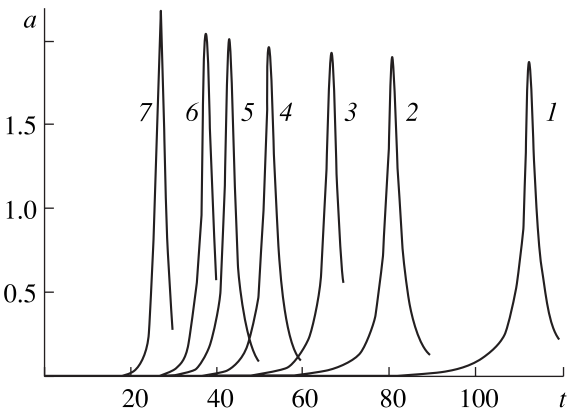


Fig. 2. Dimensionless amplitudes of nonlinearly growing waves with the dimensionless wave number $k = 1$ and dimensionless initial amplitude $a_0 = 10^{-8}$ as functions of dimensionless time (expression (12)) at $W = 2 + \Delta W$, where ΔW (the excess over the critical Tonks–Frenkel parameter) = (*1*) 0.1, (*2*) 0.2, (*3*) 0.3, (*4*) 0.5, (*5*) 0.75, (*6*) 1, and (*7*) 2.

It is worth emphasizing that the curves in Fig. 2, as well as those in Fig. 1, are meaningful only if the amplitude $a(t) < 1$. At larger values of $a(t)$, expansion (3) loses uniformity, which is central to our reasoning. Nevertheless, the presence of the maxima in the curves $a = a(t)$ allows us to qualitatively investigate the dependence of the characteristic time τ of instability development on excess W and the initial amplitude a_0 of the virtual wave. To do this, we take into account that the curves $a(t)$ sharply increase and assume that the positions of the maxima may be taken as adequate estimates of the characteristic time of instability development. We also take into account that the derivative $da(t)/dt$ vanishes at the maxima, which makes it possible to find characteristic time $\tau = \tau(a_0, \Delta W)$ as a function of the physical quantities we are interested in, namely, excess ΔW over the Tonks–Frenkel parameter and initial amplitude a_0 of the virtual wave:

$$\tau = \frac{1}{2\sqrt{\Delta W}}\ln\{-[(w_2 a_0^2 + 2\Delta W)(16\Delta W w_2 a_0^2$$
$$+ 16(\Delta W)^2 + a_0^4(w_2^2 - 12 w_4 \Delta W)) + 2\sqrt{\Delta W}$$
$$\times \sqrt{w_2 a_0^2 + \Delta W - w_4 a_0^4}(16\Delta W w_2 a_0^2$$
$$+ 16(\Delta W)^2 + a_0^4(3w_2^2 - 4w_4\Delta W))]/[2a_0^4\sqrt{\Delta W}$$
$$\times \sqrt{w_2 a_0^2 + \Delta W - w_4 a_0^4}(-w_2^2 - 4w_4\Delta W)$$
$$- a_0^4(w_2 a_0^2 + 2\Delta W)(w_2^2 + 4w_4\Delta W)]\}.$$

The plots of the function $\tau = \tau(\Delta W)$ at a_0 = const that were constructed for different ranges of ΔW using this expression are presented in Figs. 3a–3c. The dependence $\tau = \tau(a_0)$ would be more informative when calculated at $\Delta W = 0$; however, the function $\tau = \tau(a_0, \Delta W)$ obtained diverges at $\Delta W \longrightarrow 0$. Therefore, we will find the desired functional relation by equating the time derivative of amplitude $a(t)$ (given by expression (10),

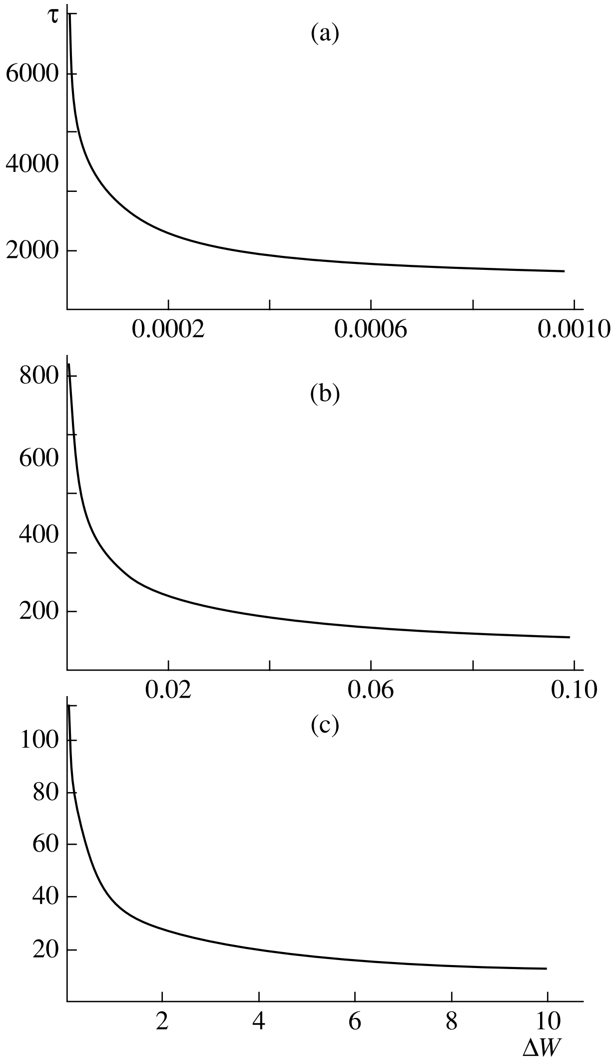


Fig. 3. Dimensionless time τ of instability development as a function of parameter ΔW for the wave with $k = 1$ and $a_0 = 10^{-8}$. The range of ΔW is (a) $10^{-5} \le \Delta W \le 10^{-3}$, (b) $10^{-3} \le \Delta W \le 10^{-1}$, and (c) $10^{-1} \le \Delta W \le 10$.

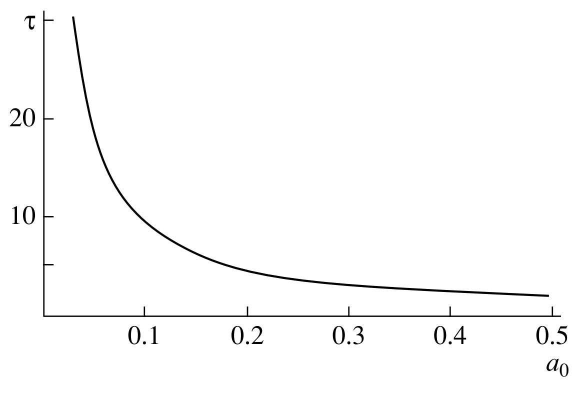


Fig. 4. Dimensionless time τ of instability development as a function of the initial wave amplitude for $k = 1$, $\Delta W = 0$, and $W = W_* = 2$.

which was derived just for the case $\Delta W = 0$) to zero. Eventually, we arrive at

$$\tau(a_0) = \frac{\sqrt{w_2 - w_4 a_0^2}}{w_2 a_0},$$

which coincides with (11). This dependence is illustrated in Fig. 4. It is easy to see that, at $W = 2$ ($\Delta W = 0$), the a_0 dependence of the characteristic time of instability development on the charged flat surface of a fluid, $\tau \approx 1/a_0\sqrt{w_2}$, differs little from the pure hyperbolic function over a wide range of the initial amplitude, $10^{-8} \le a_0 \ll 1$.

The dependences $\tau = \tau(\Delta W)$ and $\tau = \tau(a_0)$ substantiate the conclusions drawn above from indirect estimates that the characteristic time of instability development may be short (less than a second) only if the initial amplitude considerably (by four orders of magnitude) exceeds the thermal amplitude (or, in other words, if $\Delta W \ge 10^{-4}$). This circumstance should be taken into account in further experimental verifications of the critical conditions for Tonks–Frenkel instability in order to accurately measure the time of instability development—a physical quantity that has eluded the attention of researchers.

(6) Above, we have derived the time dependence of the amplitude of the unstable wave with $k = 1$ at $W \approx 2$. The fact that, according to (1), not only the wave with the wavenumber $k = 1$ but also those with $k = 2$, 3, 4, and 5 become unstable because of the nonlinear interaction calculated up to the fifth order of smallness has remained in the shade. However, it would be of interest to see how the amplitudes of the waves with $k = 2$–5 vary with time.

Without going into details, we assume (as a first approximation to a solution of the problem stated) that the dependence $a = a(t)$ derived above for the wave with $k = 1$ holds for the second- and higher order amplitude corrections to a fundamental solution. It is also assumed that the quadratic-in-a corrections to the frequencies of the waves with $k = 2$ and 3 in (1) vary in a similar manner. These assumptions are based, first, on the very classical procedure of seeking for nonlinear corrections to the fundamental wave, which was specified at the zero time [15, 21, 22] in the form of an asymptotic expansion in its amplitude a. Second, at the value of the Tonks–Frenkel parameter that is adopted in this work ($W \approx 2$), the waves with $k \ge 2$ are stable against the surface charge and their amplitudes grow only through the nonlinear interaction with the fundamental wave with $k = 1$.

In the next approximation, it should be taken into account that the field strength at the ridges of the waves increases with wave amplitude. When the amplitude of the fundamental wave with $k = 1$ becomes sufficiently large, the shorter wave with $k = 2$ appearing at the ridge of the fundamental wave may become unstable in terms of the linear analysis (i.e., criterion (2) is satisfied). The instability of the shorter wave means that the Tonks–Frenkel parameter averaged over the half-cycle of the wave with $k = 2$ exceeds $W_* = 2.5$ at the ridge of the growing wave with $k = 1$ will. As the wave amplitude rises further, the same may occur with a still shorter wave (with $k = 3$). However, such a possibility is not obvious and calls for special detailed analysis.

(7) Let us recall a result obtained from the analysis of instability of drops against self- and induced charges [17–19]. In this case, the variation of the characteristic dimensional time t_* of instability development with the amplitude ζ_0 of the initial perturbation of the equilibrium shape (the perturbation was specified in the form $\zeta_0 P_2 \cos\theta$, where $P_2\cos\theta$ is the Legendre polynomial) is shaded by the strong dependence of t_* on the radius R of the drop (the radius dependence had the form $t_* \sim R^4/\zeta_0$. At $R \sim 10^{-2}$ cm (such was the size of drops tested for stability against self-charge in the experiments [5]) and $\zeta_0 \sim 10^{-8}$ cm, we obtain $R^4/\zeta_0 \sim 1$. Therefore, the strong dependence of the characteristic time of instability development on the initial amplitude near the critical value of the Tonks–Frenkel parameter ($W = W_* = 2$), which was found in this work, is inherent to the flat charged surface of a fluid.

(8) The fact that the characteristic time of Tonks–Frenkel instability development is long (and, consequently, the velocity of fluid surface motion is very low) when the virtual wave amplitudes are infinitesimal sheds light on the influence of the fluid viscosity on the instability development mechanism. Indeed, for a most part of the characteristic time, the unstable surface moves very slowly. Accordingly, the dissipative energy losses are low and may be neglected. The fluid surface moves fast only at the final stage of instability development, the duration of which is much shorter than the total characteristic time of the process. Let us expand the time dependence of the wave amplitude (expression (10)) in powers of a_0t up to the quadratic term,

$$a(t) \approx \left\{ a_0 + a_0^2\sqrt{w_2}t + 0.5a_0^3\left[3t^2 - \frac{w_4}{w_2} - w_2\left(t^2 - \frac{w_4}{w_2^2}\right)\right]\right\}.$$

It is seen that this expansion is asymptotically correct for times $\Delta t \le (1/a_0\sqrt{w_2})$. Comparing this interval with the total time of instability development given by (11), we find that the duration of the linear stage is a good approximation of the total time of the process. The characteristic velocity of the fluid surface, $V = a_0^2\sqrt{w_2}$, is very low throughout the linear stage. Early in the nonlinear stage, when the velocity is approximated well by the term quadratic in a_0t, the acceleration of the fluid surface, $2a^2w_2$, is also minor. Consequently, the influence of the fluid viscosity on the characteristic time of instability development is weak.

Finally, the fact that the linear and quadratic (in a_0t) stages of instability development are long implies the asymptotic validity of the results (in particular, of the integral equations derived for the time evolution of the unstable wave amplitude).

CONCLUSIONS

The characteristic time of instability development on the uniformly charged flat surface of an ideal incompressible conducting fluid can be subdivided into two stages: a stage of instability initiation, or the linear stage, covering a major part of the total time and depending on the initial conditions (on the initial amplitude of the virtual wave and the excess over the critical Tonks–Frenkel parameter at the initial instant), and a very short nonlinear stage when the unstable wave amplitude increases infinitely.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 03-01-00760) and by the President of the Russian Federation (grant no. MK 929.2003.01).

REFERENCES

1. L. Tonks, Phys. Rev. **48**, 562 (1935).
2. Ya. A. Frenkel, Zh. Éksp. Teor. Fiz. **6**, 348 (1936).
3. M. D. Gabovich, Usp. Fiz. Nauk **140**, 137 (1983) [Sov. Phys. Usp. **26**, 447 (1983)].
4. J. He, N. M. Miscovsky, P. H. Cutler, and M. Chung, J. Appl. Phys. **68**, 1475 (1990).
5. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
6. G. N. De Surgy, J. P. Chabrerie, O. Denoux, and J. E. Wesfreid, J. Phys. II (France) **3**, 1201 (1993).
7. A. A. Mohamed, E. F. Elshehawey, and M. F. El-Sayed, J. Colloid Interface Sci. **169**, 65 (1995).
8. G. I. Taylor and A. D. McIwan, J. Fluid Mech. **22**, Part 1, 1 (1965).
9. J. R. Melcher, Phys. Fluids **12**, 778 (1969).
10. J. E. Allen, J. Phys. D **18**, 59 (1985).
11. A. I. Zhakin, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 3, 94 (1984).
12. A. Gonzalez and A. Castellanos, Phys. Rev. E **49**, 2935 (1994).

13. N. M. Zubarev, Zh. Éksp. Teor. Fiz. **116**, 1990 (1999) [JETP **89**, 1078 (1999)].
14. N. M. Zubarev and O. V. Zubareva, Zh. Tekh. Fiz. **71** (7), 21 (2001) [Tech. Phys. **46**, 806 (2001)].
15. D. F. Belonozhko and A. I. Grigor'ev, Zh. Tekh. Fiz. **73** (11), 37 (2003) [Tech. Phys. **48**, 1396 (2003)].
16. M. L. Aleksandrov, L. N. Gal', V. Ya. Ivanov, *et al.*, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 6, 165 (1983).
17. S. O. Shiryaeva and A. I. Grigor'ev, Zh. Tekh. Fiz. **65** (9), 39 (1995) [Tech. Phys. **40**, 889 (1995)].
18. S. O. Shiryaeva, Pis'ma Zh. Tekh. Fiz. **26** (4), 5 (2000) [Tech. Phys. Lett. **26**, 137 (2000)].
19. A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **24** (24), 36 (1998) [Tech. Phys. Lett. **24**, 962 (1998)].
20. D. F. Belonozhko and A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **29** (8), 1 (2003) [Tech. Phys. Lett. **29**, 309 (2003)].
21. A. H. Nayfeh, J. Fluid Mech. **40**, Part 4, 671 (1970).
22. A. H. Nayfeh, Phys. Fluids **13**, 545 (1970).
23. A. I. Grigor'ev, D. F. Belonozhko, and S. O. Shiryaeva, Zh. Tekh. Fiz. **69** (7), 15 (1999) [Tech. Phys. **44**, 750 (1999)].
24. A. A. Shutov, Zh. Tekh. Fiz. **72** (8), 126 (2002) [Tech. Phys. **47**, 1053 (2002)].

Translated by N. Mende

Technical Physics, Vol. 50, No. 2, 2005, pp. 167–176. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 28–36.
Original Russian Text Copyright © 2005 by Vasilyev.

GASES
AND LIQUIDS

Flow of an Electrically Conducting Bubble Liquid in the Field of an Electromagnetic Force

A. P. Vasilyev

Orenburg State University, Prospekt Pobedy 13, Orenburg, 460352 Russia
e-mail: post@mail.osu.ru

Received February 7, 2003; in final form, April 20, 2004

Abstract—Physical processes accompanying the flow of a conducting bubble liquid in crossed electric and magnetic fields are considered. Based on the general equations of mechanics of multiphase media, we develop a one-dimensional model of the flow of and heat exchange in a compressible bubble liquid when the phases are not in thermal and velocity equilibrium. The model is numerically investigated. It is demonstrated that, when the bubble liquid flows along the electromagnetic force vector, the bubbles lag behind the carrying flow and are compressed and warmed up. This causes oscillations of the bubble volume, as well as oscillations of the parameters of both the disperse and carrying phase. In particular, the compression of the bubbles reduces the volumetric gas content, as well as increases the effective conductivity of the flow and the electromagnetic force in the downstream direction. This sets conditions for crisis of the bubble flow when the electromagnetic force expels the bubbles against the main stream. On the basis of the solutions obtained, the efficiency of a gas compressor is calculated. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

Nonmachine compression of gases and vapors is of great interest in different areas of industry, since it allows for implementation of inverse thermodynamic cycles and offers possibilities of developing new-generation compressors, refrigerators, and thermal pumps for various engineering systems. In particular, it was proposed [1–3] that a gas be compressed in a magnetohydrodynamic (MHD) pump containing a low-temperature liquid metal like the gallium–indium–bismuth eutectic. For this purpose, the gas phase was injected into the liquid metal to form a bubble structure.

Gas bubbles in a stream subjected to the field of an electromagnetic force experience the action of inertial forces; expulsive forces in the pressure gradient field, which are directed opposite to the stream; and forces retarding the flow around the bubbles (drag forces) in their relative motion, which are codirected with the stream. If the drag force exceeds the expulsive electromagnetic force, the bubbles are carried by the flow toward the high pressure zone and, hence, are compressed. Otherwise, they will stop inside the channel, which results in their accumulation and makes the bubble structure unstable, or be expulsed from the magnetic field zone. The dynamic processes of the bubble flow are superposed on the processes of gas–liquid heat exchange and on the electromagnetic processes associated with a change in the effective conductivity due to a varying volumetric gas content in the flow when the bubbles are compressed.

Designing a gas compressor (including performance evaluation) is possible only by solving a system of equations of motion and heat exchange for a bubble flow.

1. STATEMENT OF THE PROBLEM

Our aim is to calculate the distribution of kinematic, dynamic, thermodynamic, and electric parameters of a monodisperse bubble flow in the adiabatic channel of a conductive MHD pump with a constant cross-sectional area ($l \times b$ = const, where l is the electrode spacing and b is the distance between the insulating walls). The pump is placed in a uniform magnetic field $\mathbf{B}(0, 0, B)$ (B = const), which abruptly vanishes at the ends of the channel of length L. It is assumed that magnetic Reynolds numbers are small, $\mathrm{Re}_m \ll 1$. Voltage U across the channel electrodes, magnetic field induction B, and the flow parameters at the entrance to the channel ($x = 0$) are assumed to be given. To solve this problem, let us invoke the equations of mechanics of a two-phase medium.

Consider the Rayleigh–Lamb system of the continuity equations for momenta and heat inflow, as well as the conditions of joint deformation of the phases in a collision-free monodisperse bubble medium where the carrying phase ($i = 1$) is a viscous incompressible conductive liquid and the disperse phase ($i = 2$) is an ideal gas [4]. Complemented with the expressions for the Ampére force and differential Ohm's law, the system has the form (phasewise)

$$\frac{\partial\rho_1}{\partial t} + \nabla^k(\rho_1 v_1^k) = 0, \quad \frac{\partial\rho_2}{\partial t} + \nabla^k(\rho_2 v_2^k) = 0, \quad \frac{\partial n}{\partial t} + \nabla^k(n v_2^k) = 0, \tag{1}$$

$$\rho_1 = \alpha_1\rho_1^{(0)}, \quad \rho_2 = \alpha_2\rho_2^0, \quad \alpha_1 + \alpha_2 = 1, \quad \frac{4}{3}\pi a^3 n = \alpha_2, \quad \rho_2^0 = p_2/(R_\mu T_2), \tag{2}$$

$$w_a = \frac{d_2 a}{dt}, \tag{3}$$

$$(1-\varphi_1)a\frac{d_2 w_a}{dt} = \frac{p_2 - p_1}{\rho_1^0} - 2\frac{\Sigma}{\rho_1^0 a} - 4\mu_1\frac{w_a}{\rho_1^0 a} - (1-\varphi_2)\frac{3}{2}w_a^2 - \frac{2}{3}\sigma_{\text{eff}}\frac{B^2 a w_a}{\rho_1^0} + (1-\varphi_3)\frac{1}{4}(\mathbf{v}_1 - \mathbf{v}_2)^2, \tag{4}$$

$$\rho_1\frac{d_1\mathbf{v}_1}{dt} = -\alpha_1\nabla P_{1*} + \alpha_1\nabla^k\tau_{1*}^k - \alpha_1 n\mathbf{f}_* + \rho_1\mathbf{g} + \mathbf{f}_1, \tag{5}$$

$$\rho_2\frac{d_2\mathbf{v}_2}{dt} = -\alpha_2\nabla P_{1*} + \alpha_2\nabla^k\tau_{1*}^k + \alpha_1 n\mathbf{f}_* + \rho_2\mathbf{g}, \tag{6}$$

$$\rho_1\frac{d_1 u_1}{dt} = -\nabla^k q_{1*}^k + nq_{\Sigma_1} + \rho_1 A_1 + q_{jv} + q_{jr}, \tag{7}$$

$$\rho_2\frac{d_2 u_2}{dt} = \frac{\alpha_2\rho_2}{\rho_2^0}\frac{d_2\rho_2^0}{dt} + nq_{\Sigma_2}. \tag{8}$$

$$\mathbf{f}_1 = \mathbf{j}_1 \times \mathbf{B}, \quad \mathbf{j}_1^0 = \sigma_1^0(\mathbf{E} + \mathbf{v}_1 \times \mathbf{B}), \tag{9, 10}$$

$$\mathbf{f}_* = \mathbf{f}_\text{m} + \mathbf{f}_\mu + \mathbf{f}_\text{e}, \quad \frac{d_i}{dt} = \frac{\partial}{dt} + (\mathbf{v}_i \cdot \nabla), \quad w_{12} = |\mathbf{v}_1 - \mathbf{v}_2|. \tag{11}$$

Here, ρ_1 and ρ_1^0 are the reduced and true densities of an ith phase, respectively; $\mathbf{v}_i$ is the velocity; n is the numerical concentration of the bubbles; a is the bubble radius; α_i is the volume content of the ith phase; T_i is the temperature; p_i is the pressure; w_a is the radial velocity of the interface; σ_{eff} is the effective conductivity of the bubble medium; P_{1*} is the spherical component of the surface stress tensor in the collision-free bubble flow; τ_{1*}^{kl} are the components of the viscous stress tensor in the carrying flow; $\mathbf{g}$ is the gravitational force density; $\mathbf{f}_1$ is the density of the electromagnetic force acting on the carrying phase in a unit volume of the mixture; $\mathbf{f}_*$ is the action of the carrying phase on the disperse one due to different velocities of the phases; $\mathbf{q}_{1*}$ is the reduced heat flux density; q_{Σ_i} is the heat flux in the ith phase at the interface; $\rho_1 A_1$ is the viscous stress power in the carrying flow in a unit volume of the mixture; q_{jv} and q_{jr} are the densities of volume heat release sources due to the conductive and eddy electric currents; $\mathbf{f}_\text{m}$ is the inertial force of apparent masses that acts on a bubble; $\mathbf{f}_\text{e}$ is the expulsive force due to the nonpotentiality of the electromagnetic force bear a bubble; $\mathbf{f}_\mu$ is the drag force (retarding the flow around a sphere); $\mathbf{j}_1$ and $\mathbf{j}_1^0$ are the reduced and true densities of the electric current, respectively; $\mathbf{E}$ is the electric field strength; and Σ is the surface tension coefficient. Functions $\varphi_1 = (1.1\alpha_2^{1/3} - \alpha_2)/\alpha_1$, $\varphi_2 = (1.5\alpha_2^{1/3} - 1.3\alpha_2)/\alpha_1$, and $\varphi_3 = \alpha_2/\alpha_1$ take into consideration the fact that the flow contains many, rather than one, bubbles [4]. Factor 2/3 in Eq. (4) appears as a result of averaging the electromagnetic force due to the eddy currents over the spherical surface [5].

We average the differential Ohm's law over the representative volume of the bubble flow, taking into account that the conductive phase occupies a fraction α_1 of the mixture volume V and that the conductivity σ_2^0 of the disperse phase is equal to zero,

$$\mathbf{j}_1 = \langle\mathbf{j}_1^0\rangle = \frac{1}{V}\int_V \mathbf{j}_1^0 dV = \frac{\alpha_1}{V_1}\int_{V_1}\sigma_1^0\mathbf{E}dV + \frac{\alpha_1}{V_1}\int_{V_1}\sigma_1^0\mathbf{v}_1 \times \mathbf{B}dV = \alpha_1\sigma_1^0\langle\mathbf{E}\rangle + \alpha_1\sigma_1^0\langle\mathbf{v}_1\rangle \times \mathbf{B} = \sigma_{\text{eff}}\left(\langle\mathbf{E}\rangle + \frac{\alpha_1}{\sigma_*}\langle\mathbf{v}_1\rangle \times \mathbf{B}\right). \tag{12}$$

Here, $\sigma_{\text{eff}} = \alpha_1\sigma_1^0 + \alpha_2\sigma_2^0$ is the effective conductivity of the bubble structure and $\sigma_* = \sigma_{\text{eff}}/\sigma_1^0$ is the effective conductivity coefficient given by [6]

$$\sigma_* = \left(1 + \frac{3}{2}\frac{(1 - \sigma_2^0/\sigma_1^0)\alpha_2}{1 - \left(1 - \dfrac{\sigma_2^0}{\sigma_1^0}\right)\sqrt[3]{\dfrac{9\pi}{16}\alpha_2^2}}\right)^{-1} = \left(1 + \frac{3}{2}\frac{\alpha_2}{1 - \sqrt[3]{\dfrac{9\pi}{16}\alpha_2^2}}\right)^{-1}_{\sigma_2^0 = 0}. \tag{13}$$

The interfacial forces due to different velocities of the phases can be written as follows [4].

The inertial force of the apparent masses is

$$\mathbf{f}_\text{m} = \frac{2\pi}{3}\rho_1^0\left(\frac{d_1\mathbf{v}_1}{dt} - \frac{d_2\mathbf{v}_2}{dt} - (\mathbf{v}_2 - \mathbf{v}_1)\frac{3}{a}\frac{d_2 a}{dt}\right), \tag{14}$$

and the drag force retarding the flow around the bubbles is

$$\mathbf{f}_\mu = \pi a^2 C_D \rho_1^0 \frac{(\mathbf{v}_1 - \mathbf{v}_2)^2}{2} \frac{\mathbf{v}_1 - \mathbf{v}_2}{|\mathbf{v}_1 - \mathbf{v}_2|}, \qquad (15)$$

where

$$C_D = C\left(1 + \frac{9}{16}\mathrm{Ha}_a\right) \; [7]$$

and

$$C = C(\mathrm{Re}_{12}) =$$

$$\begin{cases} \dfrac{24}{\mathrm{Re}_{12}} \dfrac{1 + (1/6)\sqrt[3]{\mathrm{Re}_{12}}}{1 - \alpha_2}, & \mathrm{Re}_{12} < 15 \\ \dfrac{68}{\mathrm{Re}_{12}(1 - \alpha_2)}, & 15 < \mathrm{Re}_{12} \le 500 \\ \left(\dfrac{1845}{\mathrm{Re}_{12}} + \dfrac{192}{\sqrt{\mathrm{Re}_{12}}} + 5.07\right)\dfrac{1}{1 - \alpha_2}, & \mathrm{Re}_{12} > 500 \end{cases}$$

[4, 8]. Here, $\mathrm{Re}_{12} = 2a|\mathbf{v}_1 - \mathbf{v}_2|/\nu_1^0$ and $\mathrm{Ha}_a = aB\sqrt{\sigma_1^0/\mu_1^0}$. In crossed electric and magnetic fields, a bubble experiences the action of force $\mathbf{f}_\mathrm{e}$ due to the nonuniform current spread near the sphere [9],

$$\mathbf{f}_\mathrm{e} = -\frac{m}{2}\mathbf{j}_1^0 \times \mathbf{B}\frac{4}{3}\pi a^3, \quad m = \frac{\sigma_2^0 - \sigma_1^0}{\sigma_2^0 + 2\sigma_1^0} = \frac{1}{2}\bigg|_{\sigma_2^0 = 0}.$$

It should be noted that the above system of equations is not closed, since heat fluxes q_{σ_i} at the interface, reduced heat flux density $\mathbf{q}_{1*}$, the tangential components σ_{1*}^{kl} of the stress tensor, and the dissipative function are not specified.

In the case of a one-dimensional steady flow, the missing relationships may be replaced by empirical dependences closing the system.

2. ONE-DIMENSIONAL MODEL OF A COMPRESSIBLE BUBBLE FLOW

For a steady flow, we average the above system of equations, using the methods of hydraulic theory [10] and assuming that all the parameters are uniform over the cross sections, and rearrange the equations obtained.

From the equations of phase continuity and state of gas, as well as from the law of conservation of gas mass in a bubble, reduced gas content $\alpha_* = \alpha_2(\zeta)/\alpha_2(0)$ in the flow can be expressed as

$$\alpha_*(\zeta) = \frac{v_2(0)}{v_2(x)} \frac{T_{2*}}{p_{2*}}. \qquad (16)$$

Here, $\zeta = x/L$ is the reduced longitudinal coordinate and the reduced temperature and pressure of the disperse phase are given by

$$T_{2*} = T_2(\zeta)/T_2(0), \quad p_{2*} = p_2(\zeta)/p_2(0).$$

Condition (3) of joint deformation of phases can be transformed to

$$\frac{da_*}{d\zeta} = \Pi_1 \Pi_2 \frac{1 - \alpha_2(\zeta)}{S(\zeta)} w_*(\zeta), \qquad (17)$$

where $a_* = a(\zeta)/a(0)$ is the reduced radius of a bubble. The numbers Π_1 and Π_2 of the problem, characteristic velocity w_0 for radial small-scale displacements, and phase slip coefficient S are expressed as

$$\Pi_1 = \frac{w_0}{v_1(0)}, \quad \Pi_2 = \frac{L}{\alpha_1(0) a_0},$$

$$w_0 = \sqrt{\frac{p_2(0)}{\rho_1^0}}, \quad w_* = \frac{w_a(\zeta)}{w_0}, \quad S = \frac{v_2(\zeta)}{v_1(\zeta)}.$$

In the reduced coordinates, generalized Rayleigh–Lamb equation (4) takes the form

$$\begin{aligned} \frac{dw_*}{d\zeta} = \frac{\alpha_1(\zeta)\Pi_2}{(1 - \varphi_1)S(\zeta)} \Bigg[\Pi_1 \left(p_{2*} - \frac{p_1(0)}{p_2(0)} p_{1*} \right) \\ - \frac{2}{We_1} \frac{1}{\Pi_1 a_*(\zeta)} - \frac{4}{\mathrm{Re}_1} \frac{D}{a_0} \frac{w_*(\zeta)}{a_*(\zeta)} \\ - (1 - \varphi_2)\frac{3}{2}\Pi_1 w_*^2 - \frac{2}{3}\sigma_*(\zeta) \frac{\mathrm{Ha}^2}{\mathrm{Re}_1} \frac{a_0}{D} a_*(\zeta) w_*(\zeta) \\ + \frac{1 - \varphi_3}{(1 - \alpha_2(\zeta))^2} \frac{\alpha_1^2(0)}{4\Pi_1} (1 - S(\zeta))^2 \Bigg]. \end{aligned} \qquad (18)$$

Here, the Weber number, Reynolds number, and Hartmann number are specified by the conditions at the entrance to the MHD channel,

$$\mathrm{We}_1 = \frac{\rho_1^0 v_1^2(0)}{\Sigma/a_0}, \quad \mathrm{Re}_1 = \frac{\rho_1^0 v_1(0) D}{\mu_1^0},$$

$$\mathrm{Ha} = BD\sqrt{\frac{\sigma_1^0}{\mu_1^0}}, \quad D = 4\frac{bl}{b + l},$$

and D is the hydraulic diameter of the channel.

In Eq. (8) for the heat inflow to the disperse phase, the heat fluxes from the carrying phase to the interface and from the interface to the disperse phase will be described by the Newton–Richmann heat transfer law. Heat transfer coefficients β_i involved in this law will be found through the Nusselt number. For smoothly varying flows, the dimensionless equations of heat transfer in either phase are described by the expressions [4]

$$\mathrm{Nu}_2 \approx 2, \quad \mathrm{Nu}_1 = \begin{cases} 2 + \dfrac{9}{16}\mathrm{Pe}_{1v} + \dfrac{9}{16}\mathrm{Pe}_{1v}^2, & \mathrm{Re}_{1v} < 1 \\ 2 + \dfrac{3}{5}\mathrm{Pr}_{1v}^{1/3}\mathrm{Re}_{1v}^{1/2}, & \mathrm{Re}_{1v} > 1, \end{cases}$$

where the Peclet number, Reynolds number, and Prandtl number are given by

$$\mathrm{Re}_{1v}(a_*, S) = \frac{2a_0\rho_1^0 v_1(0)\alpha_1(0)a_*(1-S)}{\mu_1^0\alpha_1},$$

$$\mathrm{Pe}_{1v}(a_*, S) = \frac{2a_0 v_1(0)\alpha_1(0)a_*(1-S)}{a_1^{(T)}},$$

$$\mathrm{Pr}_{1v} = \frac{\mathrm{Pe}_{1v}}{\mathrm{Re}_{1v}}.$$

Here, $a_1^{(T)} = \dfrac{\lambda_1^0}{\rho_1^0 c_1}$, and λ_i and c_i are the thermal conductivity and specific mass heat capacity of an ith phase, respectively.

With interphase heat exchange described in this way, energy equation (8) reduces to the equation for gas pressure in a bubble

$$\frac{dp_{2*}}{d\zeta} = 3\frac{\alpha_1(\zeta)\Pi_1\Pi_2}{a_*^2 S(\zeta)}\left(-\gamma w_*(\zeta)a_*(\zeta)\right.$$
$$\left. + \frac{1}{2}\frac{\mathrm{Nu}_2}{\Pi_1\mathrm{Pe}_a}\frac{\alpha_1(0)}{\alpha_1(\zeta)}T_{1*}\frac{1 + \dfrac{\lambda_2^0}{\lambda_1^0}\dfrac{\mathrm{Nu}_2}{\mathrm{Nu}_1}\dfrac{T_{2*}}{T_{1*}}\dfrac{T_2(0)}{T_1(0)}}{1 + \dfrac{\lambda_2^0}{\lambda_1^0}\dfrac{\mathrm{Nu}_2}{\mathrm{Nu}_1}}\right). \tag{19}$$

To rearrange energy equation (7) for the carrying phase, we assume that the channel walls are adiabatic and the heat transfer along the flow via heat conduction is much less than the convective heat transfer. On the right-hand side of Eq. (7), we take into account the viscous dissipation of energy in small-scale radial and translational displacements [4]. The work of viscous forces over a large-scale translational displacement is included in terms of the homogeneous model of bubble flow, i.e., through the hydraulic friction coefficient, with regard to the presence of a magnetic field by formulas given in [11, 12]. The Joule dissipation due to eddy currents around the bubbles is calculated by formulas given in [5], and that due to conduction currents is found under the assumption that the channel electrodes are equipotential surfaces. For such a channel, the equation of continuity for the carrying phase implies that the product $\alpha_1(\zeta)v_1(\zeta)$ is constant along the stream. Then, in a channel kept under a constant voltage (($U(\zeta)$ = const), the ratio between the electromotive force of magnetic induction, $E = \alpha_1 v_1 Bl$, and electrode voltage $U = El$ is also constant: $E/U = \alpha_1 v_1 B/E$ = const. Represent this constant in the form $(1-\theta)\sigma_*(0)$. Then, the Ohm's law in terms of the effective current density in a bubble medium is written as follows:

$$j_1(\zeta) = \sigma_{\mathrm{eff}}(\zeta)\left(E - \frac{\alpha_1 v_1 B}{\sigma_*(\zeta)}\right)$$
$$= \frac{1}{1-\theta}\left(\theta + \frac{\sigma_*(\zeta)}{\sigma_*(0)} - 1\right)\sigma_1^0\alpha_1 v_1 B.$$

It should be noted that compression of the bubbles along the channel is attended by a decrease in the volume gas content in the flow and an increase in the effective conductivity. In a channel kept at a constant voltage, this causes an increase in the electric current density and electromagnetic force density in the downstream direction.

To clarify the physical meaning of parameter θ, assume that the flow is single-phase. Then, $\alpha_1 = 1$, $\sigma_*(\zeta)/\sigma_*(0) = 1$ ($\zeta \in [0; 1]$), and the Ohm's law takes the form

$$j_1 = \frac{\theta}{1-\theta}\sigma_1^0 v_1 B.$$

In this formula, parameter θ augments the channel electric efficiency $\eta_e = v_1 Bl/U$ to unity. In what follows, parameter θ is assumed to be specified. In the above writing, the Ohm's law allows one to calculate the Joule dissipation caused by conduction currents $q_{jv} = j_1^2/\sigma_{\mathrm{eff}}$. With these remarks taken into account, energy equation (7) transforms to a temperature equation for the carrying phase,

$$\frac{dT_{1*}}{d\zeta} = \frac{\Pi_1\Pi_2}{\mathrm{Re}_a}\mathrm{Ec}_1\left(12\Pi_1^2\alpha_1\alpha_2\left(\frac{w_*}{a_*}\right)^2 + 2\Pi_1\alpha_2\left(\frac{a_0}{L}\right)^2\mathrm{Ha}^2\right.$$
$$+ \left(\frac{Q + f_*}{1-\theta}\right)^2\frac{L}{D}\frac{\mathrm{Ha}^2}{\mathrm{Re}_1}\frac{\mathrm{Re}_a}{\Pi_1\Pi_2}\frac{\alpha_1^2(0)}{\sigma_*(\zeta)}\alpha_1$$
$$+ \frac{3}{8}C_D\frac{\mathrm{Re}_a}{\Pi_1}\alpha_1^3(0)\frac{\alpha_2(1-S)^3}{a_*\alpha_1^3} + 9\frac{\alpha_1^2(0)}{\Pi_1}\frac{\alpha_2}{\alpha_1^2}\left(\frac{1-S}{a_*}\right)^2 \tag{20}$$
$$\left. - \frac{3}{2}\frac{\mathrm{Re}_a}{\mathrm{Pe}_a}\frac{\mathrm{Nu}_1}{\Pi_1\mathrm{Ec}_1}\alpha_2\frac{T_{1*}}{a_*}\left(1 - \frac{1 + \dfrac{\lambda_2^0}{\lambda_1^0}\dfrac{\mathrm{Nu}_2}{\mathrm{Nu}_1}\dfrac{T_{2*}}{T_{1*}}\dfrac{T_2(0)}{T_1(0)}}{1 + \dfrac{\lambda_2^0}{\lambda_1^0}\dfrac{\mathrm{Nu}_2}{\mathrm{Nu}_1}}\right)\right).$$

The Eckert number, Reynolds number Re_a, and Peclet number Pe_a are given by

$$\mathrm{Ec}_1 = \frac{v_1^2(0)}{c_1 T_1(0)}, \quad \mathrm{Re}_a = \frac{\rho_1^0 a_0 v_1(0)}{\mu_1^0},$$

$$\mathrm{Pe}_a = \frac{a_0 v_1(0)}{a_2^{(T)}}, \quad a_2^{(T)} = \frac{\lambda_2^0}{\rho_2^0(0) c_{2v}}.$$

The temperature T_{2*} of the disperse phase is uniquely defined by gas pressure p_{2*} in a bubble and its radius a_* from the equation of state

$$T_{2*} = p_{2*}(\zeta) a_*^3(\zeta). \tag{21}$$

To describe the motion of a bubble flow, it is more convenient to use, instead of the equations for phase momenta, the momentum equation for the whole mixture (which is obtained by adding up Eqs. (5) and (6)) and momentum equation (6) for the disperse phase. Averaging these equations by the methods of the hydraulic theory (with regard to the expression for the spherical component of the reduced stress tensor for a one-dimensional flow [4]) and neglecting the mass forces related to the gas phase density, one arrives at a system of equations for the phase slip coefficient and the pressure in the carrying flow,

$$A_1 \frac{dS}{d\zeta} + A_2 \frac{dp_{1*}}{d\zeta}$$

$$= -C_1 \frac{dw_*}{d\zeta} - C_2 \frac{dp_{2*}}{d\zeta} - C_3 + C_4 - C_5 + C_6,$$

$$B_1 \frac{dS}{d\zeta} + B_2 \frac{dp_{1*}}{d\zeta}$$

$$= D_1 \frac{dw_*}{d\zeta} - D_2 \frac{dp_{2*}}{d\zeta} - D_3 + D_4 + D_5 + D_6.$$

Denote the right-hand sides of these equations by F_1 and F_2, respectively. Assuming that the determinant of the system is nonzero, $\Delta = A_1 B_2 - A_2 B_1 \neq 0$, we solve the system for the unknown derivatives and find

$$\frac{dS}{d\zeta} = \frac{1}{\Delta}(F_1 B_2 - F_2 A_2), \tag{22}$$

$$\frac{dp_{1*}}{d\zeta} = \frac{1}{\Delta}(F_2 A_1 - F_1 B_1). \tag{23}$$

The coefficients and functions on the right-hand sides of these equations are given by the expressions

$$A_1 = -\frac{1}{2}\frac{\alpha_2}{\alpha_1}(1 - S(\zeta)) - \frac{\alpha_1(0)}{\alpha_2(0) S_0} \frac{1}{\left(1 + \frac{\alpha_1(0)}{\alpha_2(0) S_0}\frac{S}{a_*^3}\right)^2} \frac{Q_1}{a_*^3},$$

$$A_2 = \frac{\alpha_1^2(\zeta)}{\alpha_1^2(0)} \mathrm{Eu}_1, \quad C_1 = 2\frac{\alpha_2 \alpha_1}{\alpha_1^2(0)} \Pi_1^2 w_*, \quad C_2 = \frac{\alpha_2 \alpha_1}{\alpha_1^2(0)} \Pi_1^2,$$

$$C_3 = \Pi_1 \Pi_2 \frac{w_*}{S} \alpha_1$$

$$\times \left(3\frac{\alpha_1(0)}{\alpha_2(0) S_0} \frac{Q_1}{\left(1 + \frac{\alpha_1(0)}{\alpha_2(0) S_0}\frac{S}{a_*^3}\right)^2} \frac{S}{a_*^4} + \frac{2}{\mathrm{We}_1} \frac{\alpha_2 \alpha_1}{\alpha_1^2(0)} \frac{1}{a_*^2}\right),$$

$$C_4 = \frac{\theta + f_*}{1 - \theta} \frac{\alpha_1^3}{\alpha_1^2(0)} \frac{L}{D} \frac{\mathrm{Ha}^2}{\mathrm{Re}_1},$$

$$C_5 = \alpha_1 \frac{\lambda_\mathrm{m}}{2} \frac{L}{D}, \quad C_6 = \frac{\alpha_1^2}{\alpha_1^2(0)} \frac{g \delta L}{v_1^2(0)},$$

$$Q_1 = \frac{1}{\alpha_1}\left(1 + \frac{1}{2}\frac{\alpha_2}{\alpha_1}(1 - S)^2\right) + \frac{\alpha_1}{\alpha_1^2(0)}\left(\Pi_1^2 p_{2*} - Eu_1 p_{1*}\right.$$

$$\left. - \frac{2}{\mathrm{We}_1}\frac{1}{a_*} + \Pi_1^2 w_*^2\right) + \frac{1}{4}\frac{(1 - S)^2}{\alpha_1},$$

$$B_1 = \frac{1}{2}\frac{S - \alpha_2}{\alpha_1} - \frac{\alpha_1(0)}{\alpha_2(0) S_0} \frac{Q_2}{\left(1 + \frac{\alpha_1(0)}{\alpha_2(0) S_0}\frac{S}{a_*^3}\right)^2} \frac{1}{a_*^3},$$

$$B_2 = \frac{\alpha_1^2}{\alpha_1^2(0)} \mathrm{Eu}_1, \quad D_2 = \frac{\alpha_2 \alpha_1}{\alpha_1^2(0)} \Pi_1^2,$$

$$D_1 = 2\frac{\alpha_2 \alpha_1}{\alpha_1^2(0)} \Pi_1^2 w_*,$$

$$D_3 = \Pi_1 \Pi_2 \frac{w_*}{S} \alpha_1$$

$$\times \left(3\frac{\alpha_1(0)}{\alpha_2(0) S_0} \frac{Q_2}{\left(1 + \frac{\alpha_1(0)}{\alpha_2(0) S_0}\frac{S}{a_*^3}\right)^2} \frac{S}{a_*^4} + \frac{2}{\mathrm{We}_1} \frac{\alpha_1^2}{\alpha_1^2(0)} \frac{1}{a_*}\right),$$

$$D_4 = \frac{1}{4}\frac{\theta + f_*}{1 - \theta} \frac{\alpha_1}{\alpha_1(0)} \frac{L}{D} \frac{\mathrm{Ha}^2}{\mathrm{Re}_1}, \quad D_5 = \frac{3}{8} C_D \frac{L}{a_0} \frac{(1 - S)^2}{a_*},$$

$$D_6 = \frac{3}{2}\alpha_1^2 \Pi_1 \Pi_2 \frac{w_*}{a_*}(1 - S) - \frac{\lambda_\mathrm{m}}{2}\frac{L}{D},$$

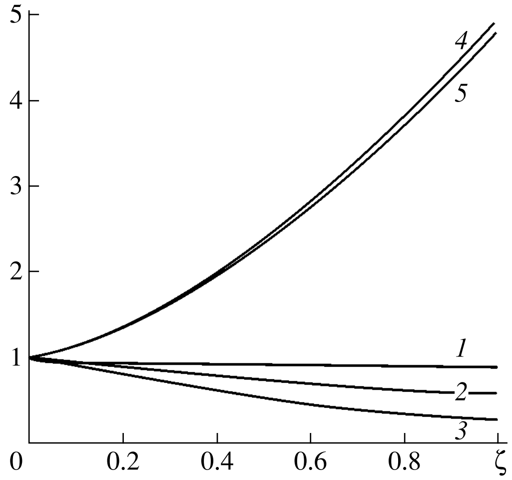


Fig. 1. Distributions of the reduced parameters of the bubble flow along the MHD channel.

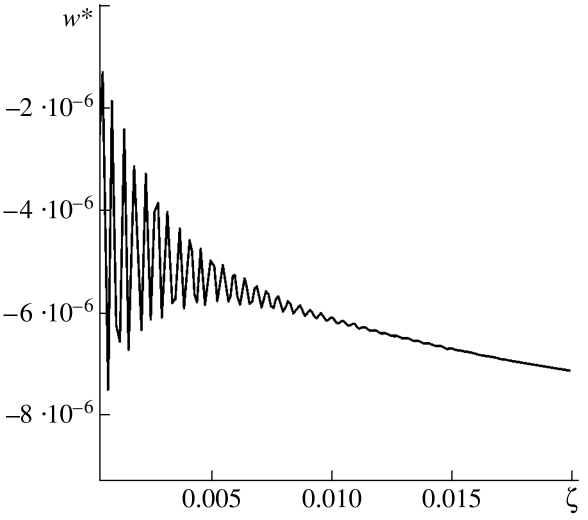


Fig. 2. Distribution of the reduced radial velocity of the interface over the initial section of the channel.

$$Q_2 = \frac{\alpha_1}{\alpha_1^2(0)}\left(\Pi_1^2 p_{2*} - \mathrm{Eu}_1 p_{1*} - \frac{2}{\mathrm{We}_1}\frac{1}{a_*} + \Pi_1^2 w_*^2\right)$$
$$+ \frac{1}{4}\frac{(1-S)^2}{\alpha_1} + \frac{1}{2\alpha_1}\left(\frac{\alpha_2}{\alpha_1}(1-S)^2 - (1-S^2)\right),$$

where $\delta = \cos(\mathbf{g}, \mathbf{i})$.

For the horizontal flow under consideration, $\delta = 0$ and the Euler number is specified by the conditions at the entrance to the channel, $\mathrm{Eu}_1 = p_1(0)/\rho_1^0 v_1^2(0)$.

Thus, we have come to a normal system of differential equations (17)–(20), (22), and (23) that is complemented by two dependences (16) and (21). The unknowns here are a_*, the bubble radius; w_*, the radial velocity; p_{1*} and p_{2*}, the phase pressures; T_{1*} and T_{2*}, the phase temperatures; S, the phase slip coefficient; and $\alpha_* = \alpha_2(\zeta)/\alpha_2(0)$, the reduced gas content.

For a bubble flow nonuniform in velocities and temperatures at the entrance to the channel, the initial conditions for the desired functions are specified in the form

$$\zeta = 0:\ S = 1,\ \ a_* = 1,\ w_* = 0,\ p_{1*} = 1,$$
$$p_{2*} = 1,\ \alpha_* = 1,\ T_{1*} = T_{2*} = 1,\ T_1(0) = T_2(0). \tag{24}$$

3. NUMERICAL INVESTIGATION OF A COMPRESSIBLE BUBBLE FLOW

The system of equations with initial conditions (24) was numerically integrated by the fourth-order Runge–Kutta method. Spatial step $\delta\zeta$ was chosen such that the results of numerical integration would remain the same even if it were half as large. Our aim was to study the effect of the flow and electromagnetic field parameters on gas compression in the bubbles. Variable parameters were magnetic field induction B, load parameter θ, the velocity $v_1(0)$ of the carrying flow at the entrance to the MHD channel, bubble initial radius a_0, and initial gas content $\alpha_2(0)$ at the entrance to the channel of fixed geometry ($l = 50$, $b = 10$ mm, $L = 1$ m). In the calculations, gallium and nitrogen were taken as the carrying phase and disperse phase, respectively.

Figures 1–3 show the solutions to the system of equations for a bubble flow in a constant-voltage channel that were obtained for the gas content at the entrance to the channel $\alpha_2(0) = 20\%$, bubble initial radius $a_0 = 0.25$ mm, magnetic field induction $B = 0.25$ T, the carrying flow velocity at the entrance $v_1(0) = 3$ m/s, and load parameter $\theta = 0.3$.

In Fig. 1, the distributions of the desired functions along the length of the MHD channel are depicted: curve *1* stands for $S(\zeta)$; curve *2*, for $a_*(\zeta)$; curve *3*, for $a_*(\zeta)$; curve *4*, for p_{1*}; and curve *5*, for p_{2*}. Figure 2 plots function $w_*(\zeta)$ for the oscillating phase of the process, and Fig. 3 plots temperature functions $T_{1*}(\zeta)$ (curve *1*) and $T_{2*}(\zeta)$ (curve *2*).

The curves in Fig.1 illustrate that the gas bubbles lag behind the carrying phase when moving in the flow. Note that the smaller the bubble radius, the closer the phase slip coefficient to unity (≈0.9 for curve *1*). The reduced radius of the bubble decreases in the downstream direction (curve *2*), which indicates that the gas in the bubble is compressed. At the entrance to the channel, the reduced gas content of the flow first increases and exceeds unity. Then, as the bubbles become compressed, it decreases and attains a value of 0.5 at the exit from the channel. The increase in the reduced gas content at the entrance of the MHD channel is due to the stagnation of the disperse phase and the

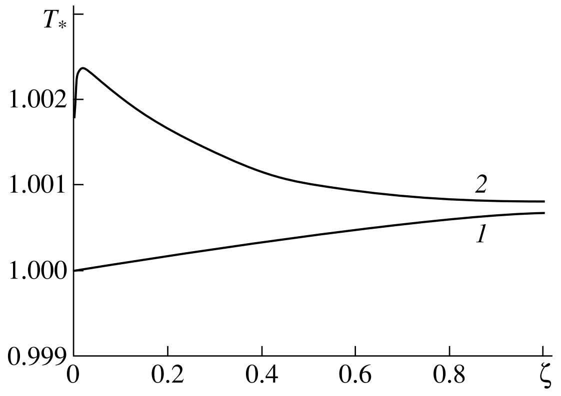


Fig. 3. Distribution of the reduced temperature in the phases along the MHD channel.

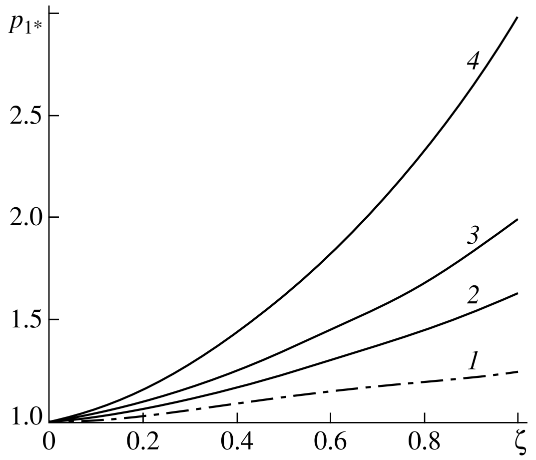


Fig. 4. Distribution of the reduced pressure in the carrying flow along the channel for various gas contents.

corresponding increase in its local concentration (curve *3*). The pressures of both the liquid and gas phases grow downstream, the liquid pressure (curve *4*) everywhere exceeding the gas pressure (curve *5*).

The difference between the pressures gives rise to a small-scale flow of the liquid around the bubble with velocity $w_*(\zeta)$ at the interface. At the entrance to the channel, this radial velocity oscillates (see Fig. 2). However, the magnetic field rapidly damps the oscillations and, as the bubbles are compressed further, the radius of the bubbles and the radial velocity vary smoothly. The carrying phase temperature (Fig. 3, curve *1*) increases along the MHD channel insignificantly, so that the carrying phase behaves like a thermostat, accumulating the heat being released during the compression of the disperse phase. The disperse phase temperature first rises (Fig. 3, curve *2*) and then decreases, while everywhere remaining higher than the liquid temperature. This decrease can be explained by more intense heat removal due to a rise in the heat transfer coefficient when the bubble radius decreases because of compression.

Figure 4 shows the variation of the carrying phase pressure with channel length ζ for various initial gas contents: curve *1* corresponds to $\alpha_2(0) = 1\%$; curve *2*, to $\alpha_2(0) = 5\%$; curve *3*, to $\alpha_2(0) = 10\%$; and curve *4*, to $\alpha_2(0) = 20\%$. The other parameters are $a_0 = 0.25$ mm, $v_1(0) = 2$ m/s, $\theta = 0.2$, and $B = 0.25$ T. At a low gas content in the flow, the dependence $p_{1*}(\zeta)$ becomes linear, which is typical of single-phase flows. As the gas content increases, the dependence $p_{1*}(\zeta)$ becomes nonlinear, which suggests that the second phase starts influencing the flow. In particular, the current density grows along the channel: the compression of the bubbles decreases the volume fraction of the gas phase in the flow, which raises the effective conductivity of the flow and, hence, the current density.

The influence of the initial radius a_0 of the bubble on the compression pattern is illustrated by Fig. 5, which shows the dependence of the bubble radius on length ζ for various a_0: curve *1* is drawn for ζ = 1 mm; curve *2*, for 0.8 mm; curve *3*, for 0.5 mm; and curve *4*, for 0.1 mm. Here, magnetic field induction $B = 0.3$ T, load parameter $\theta = 0.1$, $\alpha_2(0) = 30\%$, and $v_1(0) = 2$ m/s. As follows from these curves, smaller bubbles undergo stronger compression. On the one hand, the current density in a small-bubble flow is higher, parameter θ being the same. Indeed, small bubbles travel in the MHD channel with a larger phase slip coefficient *S*; accordingly, the local gas content is lower compared with that in a large-bubble flow and, hence, the conductivity of the small-bubble flow and the current density in it are higher. That is why a small-bubble liquid flows in the field of a high gradient of electromagnetic pressure. On the other hand, smaller bubbles undergo a higher bubble pressure $2\Sigma/a$ than larger ones. It should also be noted that, with an increase in the bubble initial radius, the dependence $a_*(\zeta)$ starts oscillating. The enlargement of this dependence for curve *2* is depicted in the inset to Fig. 5. The oscillation of the bubble volume gives rise to oscillation of other parameters of the flow, namely, the phase slip coefficient, the reduced gas content, and, as a consequence, the current density. However, a rise in the magnetic field induction effectively damps these oscillations over the initial section of the MHD channel. The flow parameters also start oscillating when the velocity of the carrying flow grows. For instance, if velocity $v_1(0)$ changes from 2 to 3 m/s for curve *3*, sustained oscillations of the parameters throughout the MHD channel are observed.

Figure 6 demonstrates the distribution of the reduced gas content along the length of the MHD channel for load parameter $\theta = 0.1$ (curve *1*), 0.2 (curve *2*), and 0.3 (curve *3*). The other parameters are $B = 0.25$ T, $\alpha_2(0) = 20\%$, $v_1(0) = 2$ m/s, and $a_0 = 0.5$ mm. An

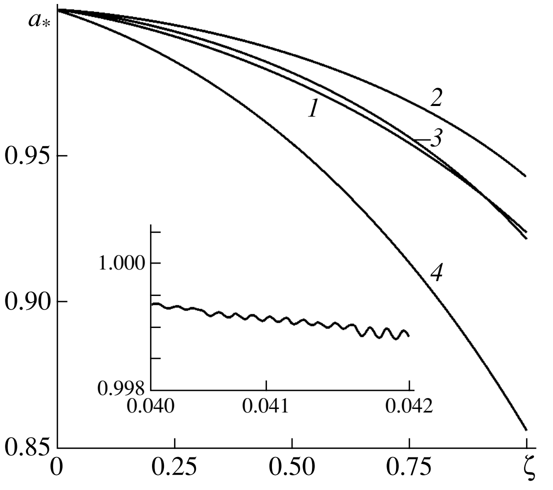


Fig. 5. Variation of the bubble reduced radius along the channel.

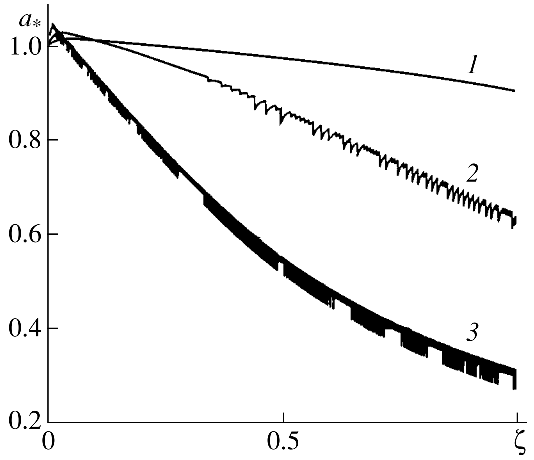


Fig. 6. Distribution of the reduced gas content along the channel for various load parameters.

increase in parameter θ (i.e., an increase in the electromagnetic force density) generates sustained oscillations of the bubbles.

Figure 7 plots the distribution of the reduced density $F_* = f_1(\zeta)/f_1(0)$ of the electromagnetic force along the MHD channel for $B = 0.3$ T, $\alpha_2(0) = 30\%$, $\theta = 0.1$, $v_1(0) = 2$ m/s, and $a_0 = 0.8$ mm. The electromagnetic force oscillations are caused primarily by the oscillations of the bubble volume.

4. ESTIMATION OF THE GAS COMPRESSION EFFICIENCY IN AN MHD COMPRESSOR

Solution of the system of equations for a bubble flow allows one to evaluate the energy performance of an MHD compressor. Indeed, from the dependence $\alpha_*(\zeta)$, one can determine the velocity of the carrying phase

$$v_1(\zeta) = v_1(0)\frac{1-\alpha_2(0)}{1-\alpha_*(\zeta)\alpha_2(0)},$$

the current density

$$j_1(\zeta) = \frac{\theta + f_*}{1-\theta}\sigma_1^0\alpha_1(\zeta)v_1(\zeta)B,$$

and the electromagnetic force density

$$f_1(\zeta) = j_1(\zeta)B, \quad f_* = \sigma_*(\zeta)/\sigma_*(0) - 1.$$

The electromagnetic force power in the active zone of the channel is

$$N_{\mathrm{EM}} = \int_{V_1}(\mathbf{j}_1^0 \times \mathbf{B}) \cdot \mathbf{v}_1 dV$$

$$= blL\frac{\theta}{1-\theta}\sigma_1(\alpha_1(0)v_1(0)B)^2\int_0^1 \frac{1+f_*/\theta}{1-\alpha_2(\zeta)}d\zeta,$$

and the electric current in the active zone of the channel is given by the integral

$$J = \int_{S_E} j_1 dS$$

$$= bL\frac{\theta}{1-\theta}\sigma_1\alpha_1(0)v_1(0)v_1(0)B\int_0^1\left(1+\frac{f_*(\zeta)}{\theta}\right)d\zeta.$$

The electric current through the end zones of the channel is calculated through the geometric conductivity C of the longitudinal effect [10],

$$J_{\mathrm{end}} = bL\frac{C}{1-\theta}\alpha_1(0)v_1(0)B\int_0^1(1+f_*(\zeta))d\zeta,$$

and the electrode voltage, through load parameter θ,

$$U = \frac{1}{1-\theta}\frac{\alpha_1(0)v_1(0)Bl}{\sigma_*(0)}.$$

The Joule dissipation in the end and active zones of an MHD channel is

$$N_{\mathrm{end}} = blL\frac{C}{(1-\theta)^2}\frac{(\alpha_1(0)v_1(0)B)^2}{\sigma_*(0)}\int_0^1(1+f_*(\zeta))d\zeta,$$

$$N_{\mathrm{J}} = blL\left(\frac{\theta}{1-\theta}\right)^2(\alpha_1(0)v_1(0)B)^2\int_0^1\frac{(1+f_*(\zeta)/\theta)^2}{1+f_*(\zeta)}d\zeta.$$

The electric power in the active zone of the channel is $N_{\mathrm{e}}^0 = N_{\mathrm{EM}} + N_{\mathrm{J}}$; the electric power supplied to the channel, $N_{\mathrm{e}} = N_{\mathrm{e}}^0 + N_{\mathrm{end}}$.

The mass flow rates of the phases are calculated from the conditions at the entrance to the channel,

$$\dot{m}_1 = \rho_1^0 \alpha_1(0) v_1(0) b l,$$

$$\dot{m}_2 = \frac{p_2(0)}{R_\mu T_2(0)} \alpha_2(0) S_0 v_1(0) b l.$$

The specific work of pressure forces in the expanded thermodynamic system (the work of compression and pushing) is given by

$$l_n = \int_{p_2(0)}^{p_2(L)} \vartheta(p)dp = R_\mu T_2(0) \int_1^{p_{2*}} T_{2*}(\zeta) \frac{dp_{2*}}{p_{2*}}$$

$$= R_\mu T_2(0) \int_0^1 T_{2*}(\zeta) \frac{dp_{2*}}{d\zeta} \frac{d\zeta}{p_{2*}(\zeta)},$$

where $\vartheta(p)$ is the specific volume of the gas and temperature T_{2*} is determined from a solution to the equation system for a bubble flow.

The useful power delivered to the gas flow by the compressor is $N_2^0 = \dot{m}_2 l_n$, and the useful power of the carrying flow is

$$N_1^0 = \dot{m}_1 \left[\frac{p_1(0)}{\rho_1^0} (p_{1*}(1) - 1) + \frac{1}{2} v_1^2(0) \left(\left(\frac{\alpha_1(0)}{\alpha_1(1)} \right)^2 - 1 - \Delta p_{\text{end}*} \right) \right],$$

where $\Delta p_{\text{end}*}$ is pressure losses in the end zones of the magnetic field due to the transverse effect that are given in terms of the dynamic pressure of the carrying flow at the entrance to the channel.

The total useful power of an MHD compressor is $N = N_1^0 + N_2^0$, and the power supply equals N_e. The ratio between these powers gives the total efficiency of an MHD compressor,

$$\eta_0 = \frac{N}{N_e} = \left(\frac{N_2^0}{N_{\text{EM}}} + \frac{N_1^0}{N_{\text{EM}}} \right) \frac{N_{\text{EM}}}{N_e^0} \frac{N_e^0}{N_e}$$

$$= (\eta_{i2} + \eta_{i1}) \eta_e \eta_{\text{end}} = \eta_i \eta_e \eta_{\text{end}}.$$

Here, factor η_i, which is equal to the ratio between the useful power and the electromagnetic power, may be called the internal efficiency. Then, η_{i1} and η_{i2} are the internal efficiencies as applied to the liquid and gas phases, respectively; η_e is the electric efficiency; and η_{end} is the "end" efficiency. The efficiency of an MHD compressor itself is $\eta_{02} = \eta_{i2}\eta_e\eta_{\text{end}}$.

For $b = 10$ mm, $l = 50$ mm, $L = 1000$ mm, $B = 0.3$ T, $\theta = 0.1$, $a_0 = 0.8$ mm, $v_1(0) = 2$ m/s, $\alpha_2(0) = 30\%$, $C = 0.022$, and $p_1(0) = 10^5$ Pa, the calculated characteristics of a constant-voltage MHD compressor were found to be as follows: the specific work $l_n = 9.8 \times 10^4$ J/kg; the nitrogen mass flow $\dot{m}_2 = 0.349$ g/s; the useful power in terms of the gas phase $N_2^0 = 34$ W; the useful power in terms of the liquid phase $N_1^0 = 12.8$ W; the total useful power of the compressor $N = 47$ W; the electric power supply $N_e = 85$ W; the internal efficiencies $\eta_i = 0.83$, $\eta_{i1} = 0.225$, and $\eta_{i2} = 0.6$; the electric efficiency $\eta_e = 0.853$; the end efficiency $\eta_{\text{end}} = 0.783$; the total efficiency in terms of the liquid phase $\eta_{01} = \eta_{i1}\eta_e\eta_{\text{end}} = 0.15$; the total efficiency in terms of the gas phase $\eta_{02} = \eta_{i2}\eta_e\eta_{\text{end}} = 0.41$; and the total efficiency of the MHD compressor $\eta_0 = \eta_i\eta_e\eta_{\text{end}} = 0.55$. Note that, with an increase in load parameter θ, the pressure ratio grows but the efficiency decreases, all other parameters of an MHD compressor being the same. For instance, when $\theta = 0.3$ and 0.4, $p_{1*}(1) = 5.83$ and 9.21, respectively.

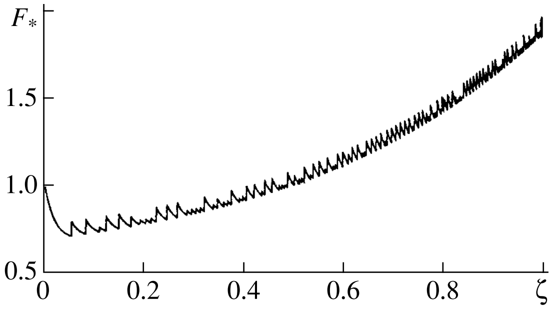


Fig. 7. Distribution of the reduced density of the electromagnetic force in the bubble flow along the MHD channel at U = const.

For each geometry and each set of the parameters of the MHD channel, there exists a limiting value $\theta_{\max}$ of the load parameter above which expulsive electromagnetic forces exceed the drag forces retarding the flow around the sphere even at the entrance to the channel. In this situation, a bubble flow becomes impossible; i.e., we are facing the problem of crisis of flow. Another type of crisis may be observed in a channel with U = const, as follows from numerical analysis: the bubbles are impeded by the increasing electromagnetic expulsion force inside the channel. As a consequence, the local gas content attains a limiting value and a bubble flow collapses.

The method of gas compression considered in this work was tested experimentally [13]. It was found that, along with the stable bubble flow, there exist critical modes of MHD compressor operation, which are associated with electromagnetic expulsion of the bubbles against the carrying stream.

CONCLUSIONS

(1) A system of equations that describes the flow of and heat exchange in a compressible bubble liquid sub-

jected to an electromagnetic force field in the channel of a conductive MHD pump is derived. Its solutions found by numerical methods using the hydraulic approach demonstrate the stability of gas compression along the stream.

(2) It is demonstrated that an increase in the initial gas content, velocity of the carrying flow, bubble initial radius, or magnetic field induction causes the parameters of a bubble flow in the MHD channel to oscillate.

(3) Under certain operating conditions, a crisis of the bubble flow arises. It is stipulated either by electromagnetic expulsion of the bubbles against the main stream or by an increase in the local gas content when the bubbles are impeded by expulsive electromagnetic forces.

(4) The total energy conversion efficiency of an MHD compressor may amount to 40–45%.

REFERENCES

1. A. P. Vasil’ev and I. M. Kirko, RF Inventor’s Certificate No. 1549445 (2000).
2. A. P. Vasil’ev and I. M. Kirko, RF Inventor’s Certificate No. 1663346 (2000).
3. A. P. Vasil’ev, Yu. A. Ponomarev, *et al.*, RF Inventor’s Certificate No. 1778459 (2000).
4. R. I. Nigmatulin, *The Fundamentals of the Mechanics of Heterogeneous Media* (Nauka, Moscow, 1978) [in Russian].
5. A. P. Vasil’ev, Zh. Tekh. Fiz. **73** (1), 35 (2003) [Tech. Phys. **48**, 31 (2003)].
6. A. P. Vasil’ev and P. V. Ivanov, Inzh.-Fiz. Zh. **47**, 693 (1984).
7. O. A. Lielausis, *Hydrodynamics of Liquid–Metal Magnetohydrodynamic Devices* (Zinatne, Riga, 1976) [in Russian].
8. A. D. Fedorovskiĭ, E. I. Nikiforovich, and N. A. Prikhod’ko, *Transfer Processes in Gas–Liquid Systems* (Naukova Dumka, Kiev, 1988) [in Russian].
9. V. I. Popov, Magn. Gidrodin., No. 3, 33 (1978).
10. A. B. Vatazhin, G. A. Lyubimov, and S. A. Regirer, *Magnetohydrodynamic Flows in Channels* (Nauka, Moscow, 1970) [in Russian].
11. A. V. Tananaev, *Ducted Flows in Magnetohydrodynamic Devices* (Atomizdat, Moscow, 1979) [in Russian].
12. *Turbulent Flows in Binary Systems*, Ed. by S. S. Kutateladze (SO AN SSSR, Novosibirsk, 1977) [in Russian].
13. A. P. Vasil’ev, V. A. Bondarenko, *et al.*, Kholodil’naya Tekhnika, No. 12, 22 (1991).

Translated by A. Pankrat’ev

Technical Physics, Vol. 50, No. 2, 2005, pp. 177–184. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 37–44.
Original Russian Text Copyright © 2005 by Belonozhko, Shiryaeva, Grigor'ev.

Nonlinear Periodic Waves on the Charged Surface of a Viscous Finite-Conductivity Fluid

D. F. Belonozhko, S. O. Shiryaeva, and A. I. Grigor'ev
Demidov State University, Sovetskaya ul. 14, Yaroslavl, 150000 Russia
e-mail: grig@uniyar.ac.ru

Received June 30, 2004

Abstract—The profile of a periodic capillary–gravitational wave propagating over the surface of a viscous finite-conductivity fluid is found in a second-order approximation in initial deformation amplitude. When the finiteness of the rate with which the potential of the fluid smoothes out as capillary–gravitational waves travel over its free surface is taken into account, the intensity of nonlinear interaction between the waves changes. This intensity is found to depend on the electric charge surface density, conductivity of the fluid, and wavenumbers. The finiteness of the potential smoothing rate influences the nonlinear interaction between the waves nonmonotonically. *© 2005 Pleiades Publishing, Inc.*

In a series of recent works concerned with the asymptotic nonlinear analysis of periodic capillary–gravitational waves on the flat surface of a deep fluid [1–4], a mathematical apparatus making it possible to correctly include the finiteness of the fluid viscosity was developed. The profile of a nonlinear periodic capillary–gravitational wave traveling on the free surface of a viscous perfectly conducting fluid and the effect of the charge uniformly distributed over the fluid surface on the wave profile were considered in detail. It was shown that second-order correction A to the initial deformation amplitude as a function of dimensionless wavenumber k (i.e., the dependence $A = A(k)$) behaves in a resonance-type manner and has a maximum near wavenumber $k_* = 1/\sqrt{2} \approx 0.707$. It is near this value of k that the nonlinear character of the wave motion shows up most vividly. The height of the peak of amplitude A, which characterizes the intensity of the internal resonant nonlinear interaction between the waves, depends on the fluid viscosity [1]: as the viscosity decreases, the peak monotonically grows and tends to infinity in the ideal fluid limit. It was demonstrated [2–4] that the intensity of the nonlinear resonant interaction is a complicated function of the surface charge density, the square of which is proportional to Tonks–Frenkel parameter W (this parameter characterizes the stability of the fluid free surface against the charge uniformly distributed over it). In the line $W = (k + k^{-1})/2$ lying on the parameter plane (k, W), amplitude A has a minimum tending to zero as the viscosity drops to zero.

With the fluid viscosity adequately taken into account, one can trace the effect of relaxation phenomena, which generate stresses tangential to the free surface, on the nonlinear interaction between capillary–gravitational waves and on nonlinear corrections to the wave profiles. In this work, we study how the finiteness of the rate of electrostatic potential smoothing over the free surface influences the nonlinear capillary–gravitational wave motion.

1. PROBLEM DEFINITION

Let a viscous incompressible fluid placed in the gravitational field occupy the semi-infinite space $z \leq 0$ and plane $0xy$ of the Cartesian system (with axis $0z$ directed oppositely to the force of gravity) coincide with the equilibrium flat free surface of the fluid. The mass density, kinematic viscosity, surface tension coefficient, and conductivity of the fluid are designated as ρ, ν, γ, and σ, respectively. We assume that the surface bears a uniformly distributed electric charge with surface density χ_0. The surface diffusion coefficient and the surface mobility of the charges are designated as D and μ, respectively, and the fluid permittivity is ε_f. Our goal is to find the profile of a planar periodic capillary–gravitational traveling wave with wavenumber $k = 2\pi/\lambda$ (λ is the wavelength) at an arbitrary time moment $t > 0$. At the zero time ($t = 0$), the wave starts propagating over the free surface of the fluid in the positive direction of axis $0x$. It is assumed that the motion of the fluid is independent of coordinate y and that the amplitude of the fundamental harmonic in the expansion of the periodic wave profile in Fourier series over spatial period λ is known and equals η. We also take into account that, as the wave propagates, the electric charge is redistributed over the deforming free surface with a characteristic time comparable to the oscillation period of the wave. In other words, the surface charge density becomes a function of time and the horizontal coordinate, $\chi = \chi(t, x)$.

Mathematically, the problem of finding the velocity field of the fluid, electrostatic potential, and profile of a nonlinear capillary–gravitational wave traveling over the charged free surface of the fluid is stated as follows:

$$\partial_t\mathbf{U} + (\mathbf{U}\cdot\boldsymbol{\nabla})\mathbf{U} = -\frac{1}{\rho}\boldsymbol{\nabla}p + \nu\Delta\mathbf{U} + \mathbf{g};\ \ \mathbf{U} = u\mathbf{e}_x + v\mathbf{e}_z;$$

$$\operatorname{div}\mathbf{U} = 0;\quad \Delta\Phi_{\text{out}} = 0;\quad \Delta\Phi_{\text{in}} = 0;$$

$$z = \xi\colon \partial_t\xi + u\partial_x\xi = v;$$

$$p - 2\rho\nu(\mathbf{n}(\mathbf{n}\cdot\boldsymbol{\nabla})\mathbf{U}) + \frac{1}{8\pi}((\boldsymbol{\nabla}\Phi_{\text{out}})^2 - \varepsilon_{\text{f}}((\mathbf{n}\cdot\boldsymbol{\nabla})\Phi_{\text{in}})^2$$

$$+ (\varepsilon_{\text{f}} - 2)((\boldsymbol{\tau}\cdot\boldsymbol{\nabla})\Phi_{\text{in}})^2) = -\frac{\gamma\partial_{xx}\xi}{(1+(\partial_x\xi)^2)^{3/2}};$$

$$-\rho\nu[(\boldsymbol{\tau}(\mathbf{n}\cdot\boldsymbol{\nabla})\mathbf{U}) + (\mathbf{n}(\boldsymbol{\tau}\cdot\boldsymbol{\nabla})\mathbf{U})] - \chi(\boldsymbol{\tau}\cdot\boldsymbol{\nabla})\Phi_{\text{in}} = 0;$$

$$\Phi_{\text{out}} = \Phi_{\text{in}};$$

$$\partial_t\chi + \sigma(\mathbf{n}\cdot\boldsymbol{\nabla})\Phi_{\text{in}} + \chi U_n\operatorname{div}_S(\mathbf{n}) + \operatorname{div}_S(\chi U_\tau\boldsymbol{\tau})$$
$$+ \mu\operatorname{div}_S(\chi E_{\text{in}_\tau}\boldsymbol{\tau}) + D\operatorname{div}_S(\operatorname{grad}_S(\chi)) = 0;$$

$$\chi = -\frac{1}{4\pi}((\mathbf{n}\cdot\boldsymbol{\nabla})\Phi_{\text{out}} - \varepsilon_{\text{f}}(\mathbf{n}\cdot\boldsymbol{\nabla})\Phi_{\text{in}});$$

$$z \longrightarrow -\infty\colon u \longrightarrow 0;\ \ v \longrightarrow 0;\ \ \boldsymbol{\nabla}\Phi_{\text{in}} \longrightarrow 0;$$

$$z \longrightarrow \infty\colon \nabla\Phi_{\text{out}} \longrightarrow -E_0\mathbf{e}_z;\quad E_0 = 4\pi\chi_0.$$

Here, Δ is the Laplacian, $\mathbf{e}_x$ and $\mathbf{e}_z$ are the respective unit vectors along the axes, and $\mathbf{n}$ and $\boldsymbol{\tau}$ are the unit vectors of the outer normal and tangent to the free surface disturbed by the wave motion of the fluid (an equation for the surface disturbed takes the form $z = \xi \equiv \xi(t, x)$; analytical expressions for $\mathbf{n}$ and $\boldsymbol{\tau}$ are given in the Appendix). Below, we will apply the technique common to the problems of nonlinear periodic waves [1–6]: the initial conditions will be determined in the course of solution so as to simplify a final result as much as possible and obtain the expression that is the most convenient for subsequent qualitative analysis of the profile.

In the statement adopted, the following functions are to be found: $\xi = \xi(t, x)$, the profile of the free surface; $u = u(t, x, z)$ and $v = v(t, x, z)$, the horizontal and vertical components of velocity field $\mathbf{U}(t, x, z)$, respectively; $p = p(t, x, z)$, the pressure distribution in the fluid; $\Phi_{\text{in}} = \Phi_{\text{in}}(t, x, z)$ and $\Phi_{\text{out}} = \Phi_{\text{out}}(t, x, z)$, the electric field potentials inside and outside the fluid, respectively; and $\chi = \chi(t, x)$, the electric charge surface density. Parameters η, k, ρ, g, ν, γ, χ_0, σ, D, μ, and ε_{f} serve as input data.

2. CONSTRUCTION OF AN ASYMPTOTIC SOLUTION

Let us take advantage of the technique described in [3, 4] to find a solution to the problem of nonlinear waves in a viscous incompressible fluid in the form

$$\begin{pmatrix} \xi \\ u \\ v \\ p \\ \Phi_{\text{out}} \\ \Phi_{\text{in}} \\ \chi \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\rho g z - \dfrac{E_0^2}{8\pi} \\ -E_0 z \\ 0 \\ \dfrac{E_0}{4\pi} \end{pmatrix} + \begin{pmatrix} \xi_1 \\ u_1 \\ v_1 \\ p_1 \\ \Phi_{\text{out}}^{(1)} \\ \Phi_{\text{in}}^{(1)} \\ \chi_1 \end{pmatrix} + \begin{pmatrix} \xi_2 \\ u_2 \\ v_2 \\ p_2 \\ \Phi_{\text{out}}^{(2)} \\ \Phi_{\text{in}}^{(2)} \\ \chi_2 \end{pmatrix} + \begin{pmatrix} O(\xi_1^3) \\ O(u_1^3) \\ O(v_1^3) \\ O(p_1^3) \\ O((\Phi_{\text{out}}^{(1)})^3) \\ O((\Phi_{\text{in}}^{(1)})^3) \\ O(\chi_1^3) \end{pmatrix};$$

$$\begin{pmatrix} \xi_2 \\ u_2 \\ v_2 \\ p_2 \\ \Phi_{\text{out}}^{(2)} \\ \Phi_{\text{in}}^{(2)} \\ \chi_2 \end{pmatrix} = \begin{pmatrix} O(\xi_1^2) \\ O(u_1^2) \\ O(v_1^2) \\ O(p_1^2) \\ O((\Phi_{\text{out}}^{(1)})^2) \\ O((\Phi_{\text{in}}^{(1)})^2) \\ O(\chi_1^2) \end{pmatrix};$$

$$\xi_1 = \eta f(t)\cos(kx - \omega t);\quad f(0) = 1.$$

Here, the quantities with sub- and superscripts 1 and 2 refer to first- and second-order amplitude corrections.

With the expansions for ξ, u, v, p, Φ_{out}, Φ_{in}, and χ, one can easily construct the first- and second-order problems (see [3, 4]):

$$\partial_t\mathbf{U}_m + \frac{1}{\rho}\boldsymbol{\nabla}p_m - \nu\Delta\mathbf{U}_m = \mathbf{V}_m;$$

$$\operatorname{div}\mathbf{U}_m = 0;\quad \Delta\Phi_{\text{out}}^{(m)} = 0;\quad \Delta\Phi_{\text{in}}^{(m)} = 0;$$

$$z = \xi\colon \partial_t\xi_m - v_m = f_{1m};$$

$$p_m - \rho g\xi_m - 2\rho\nu\partial_z v_m - \frac{E_0}{4\pi}\partial_z\Phi_{\text{out}}^{(m)} + \gamma_0\partial_{xx}\xi_m = f_{2m};$$

$$\rho\nu(\partial_z u_m + \partial_x v_m) + \frac{E_0}{4\pi}\partial_x\Phi_{\text{in}}^{(m)} = f_{3m};$$

$$\Phi_{\text{out}}^{(m)} - E_0\xi_m - \Phi_{\text{in}}^{(m)} = f_{4m};$$

$$\partial_t\chi_m + \frac{E_0}{4\pi}\partial_x u_m$$
$$+ \sigma\partial_z\Phi_{\text{in}}^{(1)} - D\partial_{xx}\chi_m - \mu\frac{E_0}{4\pi}\partial_{xx}\Phi_{\text{in}}^{(1)} = f_{5m};$$

$$\chi_m + \frac{1}{4\pi}(\partial_z\Phi_{\text{out}}^{(m)} - \varepsilon_{\text{f}}\partial_z\Phi_{\text{in}}^{(m)}) = f_{6m};$$

$$z \longrightarrow -\infty: u_m \longrightarrow 0; \quad v_m \longrightarrow 0; \quad \left|\nabla\Phi_{\text{in}}^{(m)}\right| \longrightarrow 0;$$

$$z \longrightarrow \infty: \left|\nabla\Phi_{\text{out}}^{(m)}\right| \longrightarrow 0.$$

For $m = 1$, the above relationships state the first-order problem, for which $\mathbf{V}_1 = 0$ and $f_{n1} = 0$ ($n = 1$–6). For $m = 2$, we are dealing with the second-order problem. The related quantities $\mathbf{V}_2$ and f_{n2} ($n = 1$–6) expressed through the first-order solution are given in the Appendix.

3. SOLUTION OF THE PROBLEM IN THE APPROXIMATION QUADRATIC IN PERIODIC TRAVELING WAVE AMPLITUDE

According to the routine practice in solving the problems of this class (see, e.g., [3, 4]), the next step is the solution of the problem in the first- and second-order approximations. In this way, we will obtain an expression for the profile of a periodic capillary–gravitational traveling wave in the second-order approximation in η:

$$\xi = \eta\cos\theta\exp(\delta t) + 2\eta^2[\mathrm{Re}(\zeta)\cos(2\theta) - \mathrm{Im}(\zeta)\sin(2\theta)]\exp(2\delta t); \tag{1}$$
$$\theta = \omega t - kx; \quad \omega = \mathrm{Im}(S); \quad \delta = \mathrm{Re}(S); \quad \zeta = \frac{M_1}{M_0}.$$

Here, S is the complex frequency that is derived from a dispersion relation when the first-order problem is solved and M_0 and M_1 are calculated during the solution of the second-order problem (see the Appendix).

Alternatively, expressions (1) can be written in the form

$$\xi = \eta\cos\theta\exp(\delta t) + \eta^2 A\cos(2\theta + \phi)\exp(2\delta t);$$

$$A = 2\sqrt{\mathrm{Re}(\zeta)^2 + \mathrm{Im}(\zeta)^2}; \tag{2}$$

$$\phi = \begin{cases} \arctan\left(\dfrac{\mathrm{Im}(\zeta)}{\mathrm{Re}(\zeta)}\right); & \text{if} \quad \mathrm{Re}(\zeta) > 0 \\ \dfrac{\pi}{2}; & \text{if} \quad \mathrm{Re}(\zeta) = 0 \\ \arctan\left(\dfrac{\mathrm{Im}(\zeta)}{\mathrm{Re}(\zeta)}\right) + \pi; & \text{if} \quad \mathrm{Re}(\zeta) < 0. \end{cases}$$

It should be noted that the expressions for M_0 and M_1 given in the appendix involve the reciprocal of the conductivity, $r = 1/\sigma$, instead of conductivity σ. Such a replacement simplifies asymptotic passage to the limit of a perfectly conducting fluid ($\sigma \longrightarrow \infty \Rightarrow r \longrightarrow 0$), for which the charge relaxation phenomenon is absent. In addition, as a numerical parameter characterizing the surface charge density on the flat free equilibrium surface of a fluid, we take dimensionless Tonks–Frenkel parameter [7, 8]

$$W = 4\pi\chi_0^2/\sqrt{\rho g\gamma} = E_0^2/(4\pi\sqrt{\rho g\gamma}). \tag{3}$$

Thus, quantities S, M_0, and M_1 are functions of initial parameters ρ, g, γ, ν, k, W, r, D, μ, and ε_f.

The complex frequency is calculated by the formula

$$S = \omega_0\alpha(\rho, g, \gamma, \nu, k, W, r, D, \mu, \varepsilon_f); \tag{4}$$

where

$$\omega_0^2 = kg(1 + (ak)^2 - akW)$$

(ω_0 is the frequency of infinitesimal-amplitude capillary waves with wavenumber k on the free surface of an ideal perfectly conducting fluid),

$$a = \sqrt{\gamma/\rho g}$$

is the capillary constant, and α is a dimensional root of the dimensionless dispersion relation that corresponds to a capillary–gravitational wave (for the complete dispersion relation and the choice of a proper root, see the Appendix).

It is known [8] that parameter W characterizes the stability of the uniformly charged flat surface of a fluid against its self-charge. In going to a perfectly conducting ideal fluid (ν, $r \longrightarrow 0$), the relationship for the complex frequency S takes the form $S = \pm i\omega_0$ (since $\alpha = \pm i$). Therefore, when

$$W > \frac{1}{ak} + ak \Rightarrow \omega_0^2 < 0 \Rightarrow S = \pm|\omega_0|, \tag{5}$$

that is,

$$\mathrm{Im}(S) \equiv \omega = 0, \quad \mathrm{Re}(S) \equiv \delta > 0,$$

the electric forces at the ridges of waves with wavenumber $k = 1/a$ dominate over the surface tension forces (even in the first order of smallness), making the charged surface of the fluid unstable against its self-charge [7–9]. The motion of the free surface ceases to be a wave motion, since $\omega = 0$. From (5), it readily follows that all wavenumbers $k > 0$ are stable if $0 \le W < 2$.

In view of the aforesaid, we will study the profile of wave (1) under the assumption that

$$W < \frac{1}{ak} + ak \Rightarrow \mathrm{Im}(S) = \omega \neq 0, \quad \delta = \mathrm{Re}(S) < 0. \tag{6}$$

Then, parameter δ has the meaning of the damping decrement of the wave in the first order of smallness and asymptotic approximations (1) and (2) for the wave profile remain homogeneous in the limit $\eta \longrightarrow 0$ at any time $t > 0$.

According to (2), in a quadratic approximation, the profile of a nonlinear periodic capillary–gravitational wave traveling over the free surface of a fluid is the sum of the leading (proportional to η) term (k wave) and the nonlinear-interaction-related correction, which is proportional to η^2 ($2k$ wave). Amplitude factor A in (2) serves as a measure of the intensity of interaction between these waves.

In the calculations, we used the dimensionless variables such that $\rho = g = \gamma = 1$; the other quantities were measured in terms of their characteristic scales:

$$k^* = \frac{1}{a};\quad \eta^* = a;\quad \zeta_* = \frac{1}{a};\quad r^* = \sqrt{\frac{a}{g}};$$

$$\nu^* = \sqrt{ga^3};\quad D^* = \sqrt{ga^3};\quad \mu^* = \frac{1}{\sqrt{\rho}}.$$

4. EFFECT OF CHARGE RELAXATION ON THE INTENSITY OF NONLINEAR INTERACTION BETWEEN WAVES

To study the effect of the finite conductivity of a fluid on the intensity of nonlinear interactions between the waves, we assume for simplicity that the diffusion coefficient D and the mobility μ of charge carriers equal zero (the terms proportional to them play a noticeable role only in the case of poorly conducting fluids like liquid hydrogen or helium [10–12]). This assumption implies that the electric charge relaxation in a fluid is totally governed by its conductivity. Dimensionless viscosity ν and permittivity ε_f are set equal to 0.1 and 50, respectively.

The family of curves $A = A(k)$ depicted in Fig. 1 for different r is constructed in that wavenumber range where the charge relaxation effect is the most pronounced (we recall that $A(k)$ is the amplitude factor multiplying the second-order correction to the profile of the nonlinear capillary–gravitational wave). For the given Tonks–Frenkel parameter ($W = 1$), the variation of r from 0 to 1 does not change the resonance-like shape of curves $A = A(k)$ but insignificantly (by 6%) decreases the absolute value of the peak and somewhat shifts it toward lower wavenumbers k. It is easy to check that the charge relaxation affects the nonlinear interaction intensity more appreciably in the case of short waves. It is also noteworthy that the value $r = 1$ refers to poorly conducting fluids (for which the assumption $D = \mu = 0$ is applicable with some reservations). For example, the conductivity of ethyl alcohol is $r = 1.3 \times 10^{-3}$ in terms of the dimensional variables adopted.

Yet, the surface charge density (parameter W) affects the dependence $A = A(k)$ in the case of poorly conducting fluids ($r = 1$), as well as of perfectly conducting ones, nonmonotonically (Fig. 2). As W grows from values close to zero to $W = 1$, the nonlinear inter-

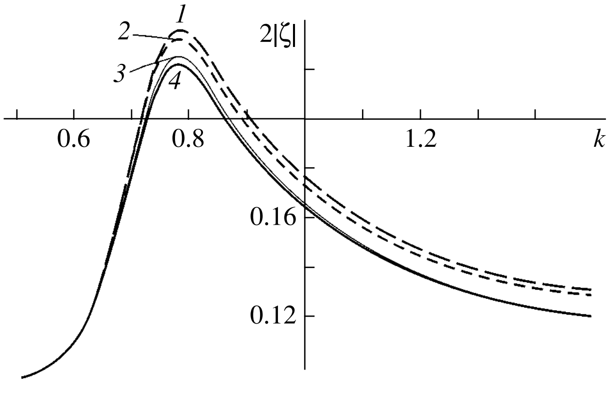


Fig. 1. Dimensionless amplitude factor $A = A(k) \equiv 2|\zeta(k)|$ multiplying the second-order correction to the wave profile vs. dimensionless wavenumber k for $W = 1$ and dimensionless fluid conductivity r = (*1*) 0, (*2*) 0.1, (*3*) 0.5, and (*4*) 1.0.

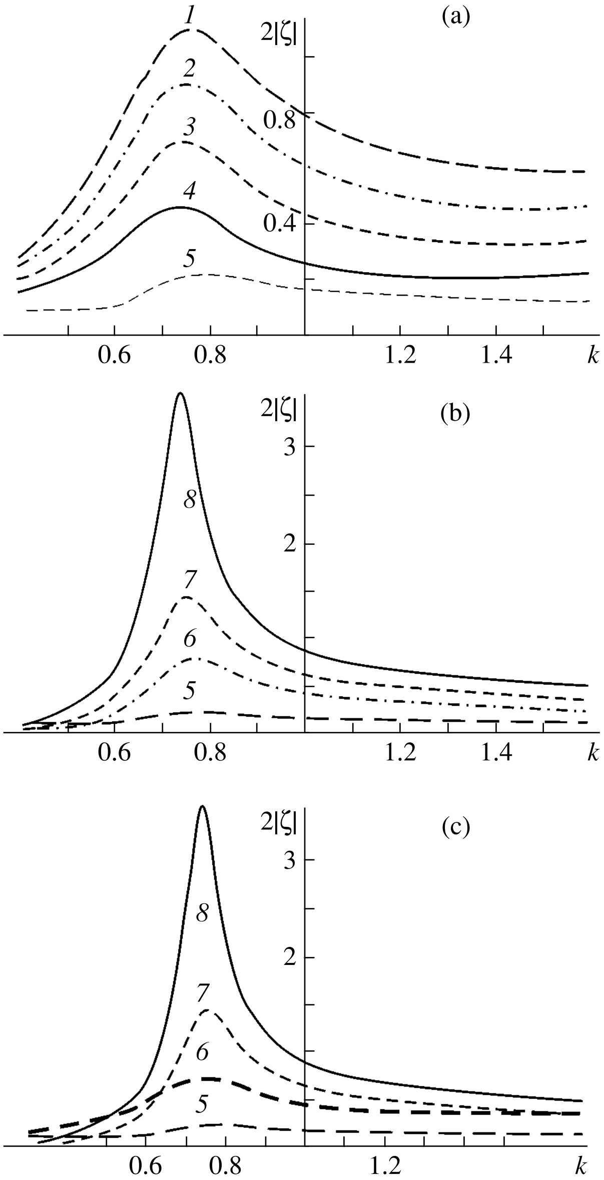


Fig. 2. Dimensionless amplitude factor $A = A(k) \equiv 2|\zeta(k)|$ multiplying the second-order correction to the wave profile vs. dimensionless wavenumber k for $r = 1$ and surface-charge-related parameter W = (*1*) 0.1, (*2*) 0.3, (*3*) 0.5, (*4*) 0.7, (*5*) 1.0, (*6*) 1.3, (*7*) 1.5, and (*8*) 1.8. (a) $W \le 1$, (b) $W \ge 1$, and (c) $0.5 \le W \le 1.8$.

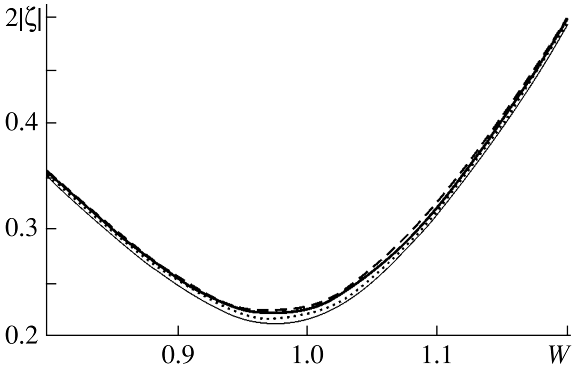


Fig. 3. Dimensionless amplitude factor $A = A(W) \equiv 2|\zeta(W)|$ multiplying the second-order correction to the wave profile vs. dimensionless parameter W, which characterizes the stability of the fluid free surface against its uniformly distributed self-charge, for $k = 0.73$ and dimensionless fluid conductivity $r = 0$ (dashed line), 0.1 (thick line),0.5 (dash-and-dot line), and 1.0 (thin line).

action intensity first declines (Fig. 2a) (the second-order correction amplitude $A(k)$ in the interval $0 \le W \le 1$ decreases) and then substantially grows in the interval $1 \le W \le 2$ (Fig. 2b). The interaction intensity is minimal at $W \approx 1$ (Fig. 2c). From Fig. 3, which plots curves $A(W)$ at different conductivities and $k = 0.73$ (at this value of k, curves $A(k)$ peak in the range where the nonlinear interaction is the most intense), it follows that this minimum for poorly conducting fluids is deeper.

Figure 4 plots amplitude A versus conductivity r for $k = 0.73$ and different values of Tonks–Frenkel parameter W. It is easy to see that this dependence noticeably depends on the surface charge (Tonks–Frenkel parameter) both qualitatively and quantitatively. It also follows from Fig. 4 that curves $A = A(r)$ run most smoothly at $W \approx 1$. That is why the peak value of amplitude $A(k)$ smoothly declines as parameter r increases (see Fig. 1). Figure 5, where a family of curves $A = A(k)$ calculated

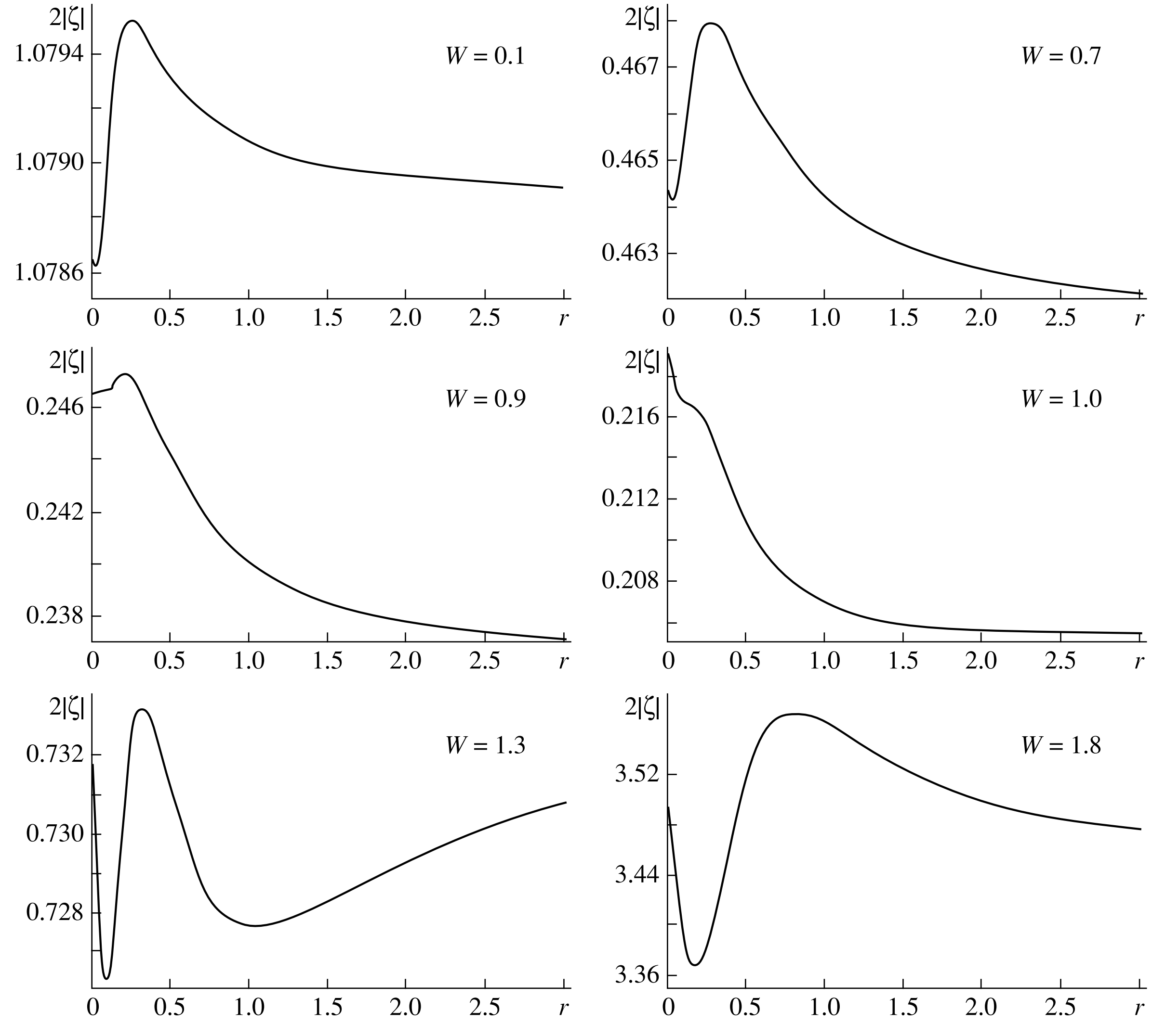


Fig. 4. Dimensionless amplitude factor $A = A(z) \equiv 2|\zeta(z)|$ multiplying the second-order correction to the wave profile vs. dimensionless fluid conductivity r for dimensionless wavenumber $k = 0.73$ and different values of W.

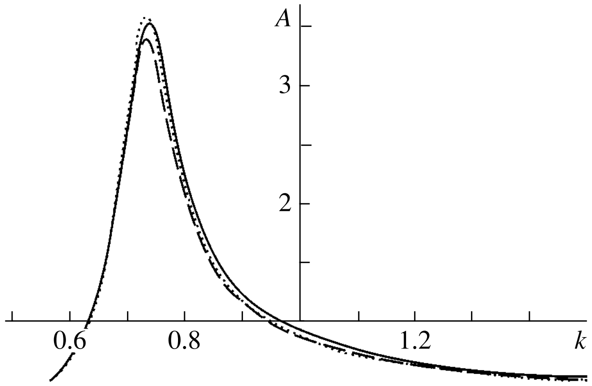


Fig. 5. Dimensionless amplitude factor $A = A(k) \equiv 2|\zeta(k)|$ multiplying the second-order correction to the wave profile vs. dimensionless wavenumber k for $W = 1.8$ and dimensionless fluid conductivity $r = 0$ (continuous line), 0.2 (dashed line), and 0.75 (dotted line).

for $W = 1.8$ is depicted, demonstrates that the monotonic decrease in nonlinear amplitude correction A with increasing r persists only for $k > 0.8$ and that the monotonicity breaks off in the range where the nonlinear mode interaction is the most intense ($k \approx k_*$). However, the absolute variation of amplitude A with magnitude of r is insignificant: in Fig. 5, the change exceeds the thickness of the lines only slightly. It should also be noted that, for $W > 1$, the interaction intensity has a dip at other-than-zero values of r (Fig. 4); i.e., this dip meets the case of finite-conductivity fluids.

CONCLUSIONS

The finiteness of the fluid conductivity most significantly influences the nonlinear interaction of capillary–gravitational waves with wavenumbers higher than $k_* = 1/\sqrt{2} \approx 0.707$ (the doubled square of this value equals unity divided by the capillary constant squared). The shape of the curves plotting the intensity of the nonlinear interaction between separate harmonics constituting a nonlinear capillary–gravitational wave versus the fluid conductivity appreciably depends on the surface charge density. As the charge density approaches a value that is critical in terms of Tonks–Frenkel instability, there appears a nonzero value of the conductivity at which the mode interaction intensity is minimal. For finite-conductivity fluids, the resonance wavenumber at which the nonlinear wave interaction is the most intense is somewhat smaller than for perfectly conducting fluids. The fluid conductivity affects the nonlinear interaction intensity to a far less extent than the surface charge: when the conductivity varies over wide limits, the interaction intensity changes within several percent, while the variation of the surface charge density in the subcritical range causes a several-fold change in the interaction intensity.

APPENDIX: AUXILIARY QUANTITIES AND RELATIONSHIPS

(i) The unit vectors tangent and normal to the perturbed free surface of the fluid:

$$\mathbf{n} = -\frac{\partial_x\xi}{\sqrt{1+(\partial_x\xi)^2}}\mathbf{e}_x + \frac{1}{\sqrt{1+(\partial_x\xi)^2}}\mathbf{e}_z;$$

$$\boldsymbol{\tau} = -\frac{1}{\sqrt{1+(\partial_x\xi)^2}}\mathbf{e}_x + \frac{\partial_x\xi}{\sqrt{1+(\partial_x\xi)^2}}\mathbf{e}_z.$$

(ii) The right-hand sides of the relationships that state the second-order problem:

$$\mathbf{V}_2 = -\frac{1}{2}\boldsymbol{\nabla}(\mathbf{U}_1^2) + \mathbf{U}_1 \times (\boldsymbol{\nabla} \times (\mathbf{U}_1));$$

$$f_{12} = \xi_1\partial_z v_1 - u_1\partial_x\xi_1;$$

$$f_{22} = 2\rho\nu\left(\xi_1\partial_{zz}v_1 + \partial_x\xi_1\left(\frac{E_0}{4\pi\rho\nu}\partial_x\Phi_{\mathrm{in}_1}\right)\right) - \xi_1\partial_z p_1$$
$$-\frac{1}{8\pi}((\boldsymbol{\nabla}\Phi_{\mathrm{out}_1})^2 - 2E_0\xi_1\partial_{zz}\Phi_{\mathrm{out}_1} - \varepsilon_{\mathrm{f}}(\partial_z\Phi_{\mathrm{in}_1})^2$$
$$+ (\varepsilon_{\mathrm{f}} - 2)(\partial_x\Phi_{\mathrm{in}_1})^2);$$

$$f_{32} = -\rho\nu(4\partial_z v_1\partial_x\xi_1 + \xi_1\partial_z(\partial_z u_1 + \partial_x v_1))$$
$$-\frac{E_0}{4\pi}\partial_x(\xi_1\partial_z\Phi_{\mathrm{in}_1}) - \chi_1\partial_x\Phi_{\mathrm{in}_1};$$

$$f_{42} = -\xi_1(\partial_z\Phi_{\mathrm{out}_1} - \partial_z\Phi_{\mathrm{in}_1});$$

$$f_{52} = -\partial_x(u_1\chi_1) - \frac{E_0}{4\pi}\left(\xi_1\partial_{xz}u_1 - \frac{E_0}{4\pi\rho\nu}\partial_x\Phi_{\mathrm{in}_1}\partial_x\xi_1\right)$$
$$-\sigma(\xi_1\partial_{zz}\Phi_{\mathrm{in}_1} - \partial_x\xi_1\partial_x\Phi_{\mathrm{in}_1}) + \mu\left(\frac{E_0}{4\pi}(\partial_{xx}\xi_1\partial_z\Phi_{\mathrm{in}_1}\right.$$
$$\left. + 2\partial_x\xi_1\partial_{xz}\Phi_{\mathrm{in}_1} + \xi_1\partial_{xxz}\Phi_{\mathrm{in}_1}) + \chi_1\partial_{xx}\Phi_{\mathrm{in}_1} + \partial_x\chi_1\partial_x\Phi_{\mathrm{in}_1}\right);$$

$$f_{62} = -\frac{1}{4\pi}\left(\xi_1\partial_{zz}(\Phi_{\mathrm{out}_1} - \varepsilon_{\mathrm{f}}\Phi_{\mathrm{in}_1})\right.$$
$$\left. - \partial_x\xi_1\partial_x(\Phi_{\mathrm{out}_1} - \varepsilon_{\mathrm{f}}\Phi_{\mathrm{in}_1}) + \frac{E_0}{2}(\partial_x\xi_1)^2\right).$$

(iii) Coefficients M_j:

$$M_j = \det$$

$$\times\begin{pmatrix} 0 & 0 & -k & ik & R_{1j} & 0 \\ \frac{kE_0}{2\pi} & 0 & -\rho(S+4\nu k^2) & 2\rho\nu ikw & R_{2j} & 0 \\ 0 & -i\frac{kE_0}{2\pi} & -4i\rho\nu k^2 & -\rho(S+4\nu k^2) & R_{3j} & 0 \\ 1 & -1 & 0 & 0 & R_{4j} & 0 \\ 4kr(2Dk^2+S) & L & -2rk^2E_0 & irkE_0w & R_{5j} & 0 \\ -\frac{k}{2\pi} & -\varepsilon_f\frac{k}{2\pi} & 0 & 0 & R_{6j} & 1 \end{pmatrix},$$

where

$$L \equiv 4k(2\pi + r(E_0\mu k + \varepsilon_f(2Dk^2+S)));$$

$$w \equiv \sqrt{2\Big(2k^2 + \frac{S}{\nu}\Big)};$$

$$R_{10} = 2S; \quad R_{20} = -(\rho g + 4\gamma k^2); \quad R_{30} = 0;$$

$$R_{40} = -E_0; \quad R_{50} = 0; \quad R_{60} = 0;$$

$$R_{11}$$

$$= \frac{1}{2}k\Big(bk\Big(1 - \frac{ic(k^2-q^2)}{(3k+q)(2S+\nu(k-q)(3k+q))}\Big) - icq\Big);$$

$$R_{21} = \frac{1}{4}k\Big(bS\rho + 2\rho\nu(bk^2 - icq^2)$$

$$+ \frac{E_0^2k}{4\pi}(a_{\text{out}} + ((\varepsilon_f - 1)a_{\text{in}} - 2)a_{\text{in}})$$

$$- \frac{2\rho ibck(k-q)(2S+\nu(5k^2+2kq+q^2))}{(3k+q)(2S+\nu(k-q)(3k+q))}\Big);$$

$$R_{31} = \frac{1}{4}\Big(\rho\nu\Big(cq(5k^2+q^2)$$

$$+ bk\Big(6ik^2 + \frac{c(k^2-q^2)(5k^2+2kq+q^2)}{(3k+q)(2S+\nu(k-q)(3k+q))}\Big)\Big)$$

$$+ i\frac{E_0^2k^2}{4\pi}a_{\text{in}}(2 + a_{\text{out}} + \varepsilon_f a_{\text{in}})\Big);$$

$$R_{41} = k\frac{E_0}{4}(a_{\text{out}} + a_{\text{in}});$$

$$R_{51} = \frac{r}{4}kE_0\Big(-icq(q + 2k(a_{\text{out}} + \varepsilon_f a_{\text{in}}))$$

$$+ bk\Big(k(1 + 2(a_{\text{out}} + \varepsilon_f a_{\text{in}}))$$

$$- \frac{2ic(k-q)(k+q)^2}{(3k+q)(2S+\nu(k-q)(3k+q))}\Big)$$

$$- k\Big((2Dk^2+S)(1 - 4(a_{\text{out}} - \varepsilon_f a_{\text{in}}))$$

$$+ a_{\text{in}}\Big(\frac{E_0^2}{4\pi\nu\rho} + 2E_0\mu k(2 + a_{\text{out}} + \varepsilon_f a_{\text{in}})\Big)\Big)\Big) + 2\pi E_0k^2a_{\text{in}};$$

$$R_{61} = \frac{E_0k^2}{32\pi}(1 - 4(a_{\text{out}} - \varepsilon_f a_{\text{in}}));$$

$$b = -\frac{\omega_0^2 + 2\nu kqS + \frac{r}{4\pi G}\frac{E_0^2k}{4\pi\rho}((k-q)S + k^3D)}{k\Big(S + 2\nu k(k-q) - \frac{r}{4\pi G}\frac{E_0^2k}{4\pi\rho}(k-q)\Big)};$$

$$c = \frac{i\Big(\omega_0^2 + S(S+2\nu k^2) + \frac{r}{4\pi G}\frac{E_0^2k}{4\pi\rho}k^3D\Big)}{k\Big(S + 2\nu k(k-q) - \frac{r}{4\pi G}\frac{E_0^2k}{4\pi\rho}(k-q)\Big)};$$

$$G = 1 + \frac{r}{4\pi}((1+\varepsilon_f)(S + k^2D) + E_0k\mu);$$

$$d = \frac{E_0k}{4\pi}\Bigg(1 - (1+\varepsilon_f)\frac{r}{4\pi G}$$

$$\times\frac{(k-q)(\omega_0^2 + S^2 + 2\nu k^2(S + k^2D)) + Sk^3D}{k\Big(S + 2\nu k(k-q) - \frac{r}{4\pi G}\frac{E_0^2k}{4\pi\rho}(k-q)\Big)}\Bigg);$$

$$a_{\text{out}} = 1$$

$$- \frac{r}{4\pi G}\frac{(k-q)(\omega_0^2 + S^2 + 2\nu k^2(S + k^2D)) + Sk^3D}{k\Big(S + 2\nu k(k-q) - \frac{r}{4\pi G}\frac{E_0^2k}{4\pi\rho}(k-q)\Big)};$$

$$a_{\text{in}} = -\frac{r}{4\pi G}$$

$$\times\frac{(k-q)(\omega_0^2 + S^2 + 2\nu k^2(S + k^2D)) + Sk^3D}{k\Big(S + 2\nu k(k-q) - \frac{r}{4\pi G}\frac{E_0^2k}{4\pi\rho}(k-q)\Big)};$$

$$q = \sqrt{k^2 + \frac{S}{\nu}};$$

$$\omega_0^2 = kg(1 + (ak)^2 - akW); \quad a = \sqrt{\frac{\gamma}{\rho g}};$$

$$E_0 = 2\sqrt{\pi W \sqrt{\rho g \gamma}},$$

and i is the imaginary unit.

(iv) The dimensionless dispersion relation has the form

$$\begin{cases} F(\alpha, \beta, R, E, \Delta, M, \varepsilon_f) = \sqrt{\alpha + \beta^2}; \\ \mathrm{Re}(F(\alpha, \beta, R, E, \Delta, M, \varepsilon_f)) > 0, \end{cases}$$

where

$$F(\alpha, \beta, R, E, \Delta, M, \varepsilon_f)$$

$$= \frac{\left((\alpha + 2\beta^2)^2 + 1 - \dfrac{R(\alpha^2 + 1 + 2\beta^2(2\alpha + \Delta))E^2}{\alpha(1 + R((1 + \varepsilon_f)(\alpha + \Delta) + \mathrm{EM}))}\right)}{4\beta^3\left(1 - \dfrac{R(\alpha^2 + 1 + 2\beta^2(2\alpha + \Delta))E^2}{\alpha(1 + R((1 + \varepsilon_f)(\alpha + \Delta) + \mathrm{EM}))4\beta^4}\right)};$$

$$\beta = \sqrt{\frac{\nu k^2}{\omega_0}}; \quad R = \frac{r\omega_0}{4\pi};$$

$$\Delta = \frac{Dk^2}{\omega_0}; \quad E = \frac{E_0 k}{\omega_0\sqrt{4\pi\rho}}; \quad M = \mu\sqrt{4\pi\rho}.$$

In the general case, the dispersion relation has two pairs of complex conjugate roots. To avoid awkwardness, we will consider only waves propagating in the positive direction of axis 0x. To this end, we take the roots with positive imaginary parts. One such root corresponds to a capillary gravitational wave, while the other, to a wave due to the relaxation of the surface electric charge.

Let these roots be known for given $\beta = \beta_*$, $E = E_*$, $\Delta = \Delta_*$, $M = M_*$, and $R = R_*$. If we put $\beta = \beta_*$, $E = E_*$, $\Delta = \Delta_*$, $M = M_*$, and $R = 0$ in the dispersion relation (i.e., if we consider a perfectly conducting fluid, for which the relaxation wave is absent), there will be only one root with a positive imaginary part. Then, varying R from 0 to R_*, we trace the variation of the root found (the procedure is accomplished numerically). At $R = R_*$, this root will equal to one of those found previously and, thus, correspond to a capillary–gravitational wave.

ACKNOWLEDGMENTS

This work was supported by the President of the Russian Federation (grant no. MK-929.2003.01) and the Russian Foundation for Basic Research (grant no. 03-01-00760).

REFERENCES

1. D. F. Belonozhko and A. I. Grigor’ev, Zh. Tekh. Fiz. **73** (4), 28 (2003) [Tech. Phys. **48**, 404 (2003)].
2. D. F. Belonozhko and A. I. Grigor’ev, Pis’ma Zh. Tekh. Fiz. **29** (8), 1 (2003) [Tech. Phys. Lett. **29**, 309 (2003)].
3. D. F. Belonozhko and A. I. Grigor’ev, Zh. Tekh. Fiz. **73** (11), 37 (2003) [Tech. Phys. **48**, 1396 (2003)].
4. D. F. Belonozhko and A. I. Grigor’ev, Zh. Tekh. Fiz. **74** (3), 5 (2004) [Tech. Phys. **49**, 287 (2004)].
5. W. F. Simmons, Proc. R. Soc. London, Ser. A **309**, 551 (1969).
6. A. H. Nayfeh, J. Fluid Mech. **48**, Part 2, 385 (1971).
7. D. F. Belonozhko, S. O. Shiryaeva, and A. I. Grigor’ev, Pis’ma Zh. Tekh. Fiz. **22** (15), 61 (1996) [Tech. Phys. Lett. **22**, 626 (1996)].
8. Ya. I. Frenkel, Zh. Éksp. Teor. Fiz. **6**, 348 (1936).
9. S. O. Shiryaeva, D. F. Belonozhko, and A. I. Grigor’ev, Zh. Tekh. Fiz. **68** (2), 22 (1998) [Tech. Phys. **43**, 151 (1998)].
10. L. P. Gor’kov and D. M. Chernikova, Dokl. Akad. Nauk SSSR **228**, 829 (1976) [Sov. Phys. Dokl. **21**, 328 (1976)].
11. A. P. Volodin, M. S. Khaĭkin, and V. S. Édel’man, Pis’ma Zh. Éksp. Teor. Fiz. **26**, 707 (1977) [JETP Lett. **26**, 543 (1977)].
12. N. M. Zubarev, Zh. Éksp. Teor. Fiz. **121**, 624 (2002) [JETP **94**, 534 (2002)].

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 185–192. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 45–52.
Original Russian Text Copyright © 2005 by Shiryaeva, Belonozhko, Grigor'ev.

GASES AND LIQUIDS

On the Internal Nonlinear Resonant Three-Mode Interaction of Charged Drop Oscillations

S. O. Shiryaeva, D. F. Belonozhko, and A. I. Grigor'ev
Demidov State University, Sovetskaya ul. 14, Yaroslavl, 150000 Russia
e-mail: grig@uniyar.ac.ru

Received June 30, 2004

Abstract—The difference between internal nonlinear three-mode degenerate and Raman resonances is found for the first time: in the former case, the energy spent on the initial deformation of a drop is only transferred from lower to higher modes; in the latter case, it is transferred in both directions. It turns out that degenerate resonances are slightly sensitive to the physical quantities that are responsible for the exact positions of the resonances (i.e., to the amount of electric charge). A deviation from the resonant value only changes the fraction of the energy the modes exchange and the time of resonant energy exchange: the interaction itself remains resonant. © *2005 Pleiades Publishing, Inc.*

(1) Among the effects related to nonlinear oscillations of a charged conducting incompressible liquid drop, internal nonlinear resonant interaction between oscillation modes occupies a prominent place, to judge by the number of publications concerned with this issue. From the pioneering works [1–5], appearing 20 years ago, to this day [6–18], more than three-fourths of the articles touching upon the problem of nonlinear oscillations of a charged drop have been devoted to the internal resonant interaction in one way or another. The fact is that resonant interaction provides the fastest and most efficient redistribution of the initial deformation energy of the drop between the modes excited via nonlinear interaction. To put otherwise, resonant interaction exerts a decisive effect both on the development of nonlinear oscillations (and, accordingly, the generation of their related acoustic and electromagnetic radiations [12, 14]) and on the disintegration of the drop that bears a near-critical (in terms of linear stability) charge [2, 5, 9, 11, 15, 17]. In spite of a great deal of articles devoted to resonant mode interaction, many related issues remain to be understood, the direction of resonant energy transfer between modes among them. Degenerate three-mode resonances, where one of two modes interacts with the other twice, were discovered and studied first [1–3]. It was argued [10, 16] that the energy in such resonances is transferred only from lower to higher modes, which, generally speaking, is inconsistent with the concept of disintegration instability as applied to three-mode interactions [19]. Furthermore, it was found [13] that disintegration instability may occur at truly three-mode resonances (secondary Raman resonances), specifically, that there are a number of resonant situations when the energy is transferred from two higher modes to a third (lower) one. However, the parameters of such an interaction (the time and amount of interaction) have not been explored. In [17], where four-mode interactions were studied, it was also demonstrated that energy may be transferred from higher to lower modes but the intensity of such transfer is low, since these interactions have the third order of smallness. Energy transfer from higher modes of nonlinear oscillations to lower ones (more specifically, to the fundamental mode) is of special interest in view of the mechanism of corona-initiated lightning discharge near a coarse highly charged drop, which is being discussed in the literature [15, 18].

In this work, we perform a detailed study of the energy transfer between modes that occurs in degenerate and secondary Raman resonances under three-mode interaction.

(2) Consider the time evolution of the surface of a nonlinearly oscillating drop of an ideal incompressible conducting liquid. The drop has a radius R, density ρ, surface tension coefficient γ, and charge Q that is uniformly distributed over the surface. At the zero time $t = 0$, the equilibrium shape of the drop experiences an axisymmetric perturbation of fixed amplitude that is much smaller than the radius of the drop. Our aim is to find the spectrum of oscillations of the drop at $t > 0$.

We assume that the drop is axisymmetric from the initial time on, so that the equation of its surface in the spherical coordinate system with the origin placed at the center of the drop has the form

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

(we use dimensional variables such that $\rho = R = \gamma = 1$).

The flow of the liquid in the drop is assumed to be potential with a velocity field potential $\psi(\mathbf{r}, t)$. Velocity field $\mathbf{V}(\mathbf{r}, t)$, in turn, is specified by a potential gradient: $\mathbf{V}(\mathbf{r}, t) = \mathrm{grad}(\psi(\mathbf{r}, t))$. If the hydrodynamic velocity of the liquid in the drop is taken to be much lower than the propagation velocity of electromagnetic interactions, the electric field of charge Q near the drop can be assumed to be electrostatic. Then, it can be described by potential $\mathbf{\Phi}(\mathbf{r}, t)$, which is related to field strength $\mathbf{E}$ as $\mathbf{E} = -\mathrm{grad}(\Phi)$.

Mathematically, the problem is stated as

$$\Delta\psi(\mathbf{r}, t) = 0; \quad \Delta\Phi(\mathbf{r}, t) = 0; \qquad (2)$$

$$r \longrightarrow 0: \psi(\mathbf{r}, t) \longrightarrow 0; \qquad (3)$$

$$r \longrightarrow \infty: |\mathrm{grad}(\Phi(\mathbf{r}, t))| \longrightarrow 0; \qquad (4)$$

$$r = 1 + \xi(\theta, t): \frac{\partial\xi}{\partial t} = \frac{\partial\psi}{\partial r} - \frac{1}{r^2}\frac{\partial\psi}{\partial\theta}; \qquad (5)$$

$$\Delta p - \frac{\partial\psi}{\partial t} - \frac{1}{2}(\boldsymbol{\nabla}\psi)^2 + \frac{1}{8\pi}(\boldsymbol{\nabla}\Phi)^2 = \mathrm{div}\,\mathbf{n}; \qquad (6)$$

$$\Phi(r, \theta, t) = \mathrm{const}; \qquad (7)$$

$$\int_V r^2 dr \sin\theta d\theta d\varphi = \frac{4}{3\pi}, \qquad (8)$$
$$V = [0 \le r \le 1 + \xi(\theta, t),\ 0 \le \theta \le \pi,\ 0 \le \varphi \le 2\pi];$$

$$\int_V \mathbf{e}_{\mathrm{r}} \cdot r^3 dr \sin\theta d\theta d\varphi = 0; \qquad (9)$$

$$-\frac{1}{4\pi}\oint_S (\mathbf{n} \cdot \boldsymbol{\nabla}\Phi) ds = Q, \qquad (10)$$
$$S = [r = 1 + \xi(\theta, t),\ 0 \le \theta \le \pi,\ 0 \le \varphi \le 2\pi];$$

$$t = 0: \xi(\theta) = \xi_0 P_0(\mu) + \xi_1 P_1(\mu) + \varepsilon \sum_{i \in \Xi} h_i P_i(\mu); \qquad (11)$$
$$\sum_{i \in \Xi} h_i = 1; \quad \frac{\partial\xi(\theta, t)}{\partial t} = 0.$$

Here, Δ is the Laplacian.

Since conditions (8) and (9) must be fulfilled at any, including initial, time instant, they define (at $t = 0$) the amplitudes of the zeroth and first modes in the expansion of the equilibrium (spherical) shape of the drop, $\xi(\theta)$, in Legendre polynomials. This means that the amplitudes of both modes cannot be taken arbitrarily: they will depend on the initial deformation.

In expressions (6)–(11), $\mu = \cos\theta$; Δp is the difference in the pressures inside and outside the drop in equilibrium; $\mathbf{n}$ is the unit normal vector to surface (1); ε is the amplitude of a small initial perturbation of the surface (the small parameter of the problem); $P_i(\mu)$ are the ith-order Legendre polynomials; h_i are the coefficients specifying the partial contribution of an ith vibrational mode to the total initial perturbation; Ξ is a set of the numbers of initially excited vibrational modes; and

$$\xi_0 \approx -\varepsilon^2 \sum_{m=1}^{\infty} \frac{h_i^2}{(2i+1)} + O(\varepsilon^3);$$
$$\xi_1 \approx -\varepsilon^2 \sum_{i \in \Xi} \frac{9 i h_{i-1} h_i}{(2i-1)(2i+1)} + O(\varepsilon^3) \qquad (12)$$

are the constants that are found from conditions (8) and (9) at the zero time (in (12), they are given accurate to the third order of smallness in ε).

(3) To find a solution to the problem stated, we will take advantage of the method of many scales (used to solve similar problems considered in [2, 5–7, 9–18]). Desired functions $\xi(\theta, t)$, $\psi(\mathbf{r}, t)$, and $\Phi(\mathbf{r}, t)$ are represented as series in powers of small parameter ε and are assumed to be independent not merely of time t but of various ε-defined times $T_m \equiv \varepsilon^m t$:

$$\xi(\theta, t) = \sum_{m=1}^{\infty} \varepsilon^m \xi^{(m)}(\theta, T_0, T_1, \ldots);$$
$$\psi(\mathbf{r}, t) = \sum_{m=1}^{\infty} \varepsilon^m \psi^{(m)}(r, \theta, T_0, T_1, \ldots); \qquad (13)$$
$$\Phi(\mathbf{r}, t) = \sum_{m=0}^{\infty} \varepsilon^m \Phi^{(m)}(r, \theta, T_0, T_1, \ldots).$$

We will restrict our analysis to a quadratic approximation and seek for dependences of the desired quantities on time scales T_0 and T_1.

Substituting expansions (13) into set (2)–(11) and equating the terms of the same power of ε, we arrive at a set of boundary-value problems for functions $\xi^{(m)}$, $\psi^{(m)}$, and $\Phi^{(m)}$. Obviously, each of functions $\psi^{(m)}$ and $\Phi^{(m)}$ must satisfy linear equations (2).

In the zeroth order of smallness, we obtain expressions for the electrostatic potential near an equilibrium (spherical) drop with charge Q, $\Phi^{(0)} = Q/r$.

The first- and second-order solutions to Eqs. (2) that satisfy boundedness conditions (3) and (4) are written in the form

$$\psi^{(m)}(r, \theta, T_0, T_1) = \sum_{n=1}^{\infty} D_n^{(m)}(T_0, T_1) r^n P_n(\mu) \ (m = 1, 2);$$
$$\Phi^{(m)}(r, \theta, T_0, T_1) = \sum_{n=0}^{\infty} F_n^{(m)}(T_0, T_1) r^{-(n+1)} P_n(\mu). \qquad (14)$$

First- and second-order corrections to the equilibrium surface of the drop are also represented as expansions in Legendre polynomials:

$$\xi^{(m)}(\theta, T_0, T_1) = \sum_{n=0}^{\infty} M_n^{(m)}(T_0, T_1)P_n(\mu) \quad (m = 1, 2). \tag{15}$$

Substituting solutions (14) and (15) at $m = 1$ into the set of first-order boundary conditions that is derived from (5)–(7) and carrying out necessary transformations, we arrive at differential equations for coefficients $M_n^{(1)}(T_0, T_1)$:

$$\frac{\partial M_n^{(1)}(T_0, T_1)}{\partial T_0^2} + \omega_n^2 M_n^{(1)}(T_0, T_1) = 0;$$
$$\omega_n^2 = n(n-1)((n+2)-W); \quad W = \frac{Q^2}{4\pi}. \tag{16}$$

A solution to (16) is a set of harmonic functions with T_1-dependent coefficients:

$$M_n^{(1)}(T_0, T_1) = A_n^{(1)}(T_1)\exp(i\omega_n T_0) + \text{c.c.},$$
$$A_n^{(1)}(T_1) = a_n^{(1)}(T_1)\exp(ib_n^{(1)}(T_1)) \quad (n \ge 2), \tag{17}$$

where c.c. hereafter stands for complex conjugate and $a_n^{(1)}(T_1)$ and $b_n^{(1)}(T_1)$ are real functions for which the dependence on T_1 can be found by only solving the problem in the next order of smallness.

From conditions (9) and (10), which are written in the approximation linear in small parameter ε, it follows that

$$M_0^{(1)}(T_0, T_1) = 0; \quad M_1^{(1)}(T_0, T_1) = 0. \tag{18}$$

Note that, formally, expressions (18) do not contradict Eqs. (16) at $n = 0$ and 1.

Satisfying initial conditions (1) in the first-order approximation in ε, we get

$$a_i^{(1)}(0) = \frac{1}{2}h_i; \quad b_i^{(1)}(0) = 0 \quad (i \in \Xi);$$
$$a_n^{(1)}(0) = 0; \quad b_n^{(1)}(n) = 0 \quad (n \notin \Xi). \tag{19}$$

Now we substitute first-order solutions (17) and (18), as well as solutions (14) and (15) at $m = 2$, into the set of second-order boundary conditions (that has been found from (5)–(7)) and, after tedious transformations, come to an equation for unknown coefficients $M_n^{(2)}(T_0, T_1)$:

$$\frac{\partial M_n^{(2)}(T_0, T_1)}{\partial T_0^2} + \omega_n^2 M_n^{(1)}(T_0, T_1) = -2i\omega_n \frac{dA_n^{(1)}(T_1)}{dT_1}$$
$$\times \exp(i\omega_n T_0) + \sum_{l=2}^{\infty}\sum_{m=2}^{\infty}\{(\gamma_{lmn} + \omega_l\omega_m\eta_{lmn})$$
$$\times A_l^{(1)}(T_1)A_m^{(1)}(T_1)\exp(i(\omega_l + \omega_m)T_0)$$
$$+ (\gamma_{lmn} - \omega_l\omega_m\eta_{lmn})A_l^{(1)}(T_1)\overline{A_m^{(1)}(T_1)}$$
$$\times \exp(i(\omega_l - \omega_m)T_0) + \text{c.c.}\};$$

$$\gamma_{ijn} = K_{ijn}\Big[\omega_i^2(n-i+1) + 2n(j(j+1)-1) + (j(i+1)$$
$$- i(2i - 2n + 7) + 3)n\frac{W}{2}\Big] + \alpha_{ijn}\Big[\frac{1}{i}\omega_i^2 + n\frac{W}{2}\Big];$$
$$\eta_{ijn} = K_{ijn}\Big(\frac{n}{2} - i + 1\Big) + \alpha_{ijn}\frac{1}{i}\Big(1 + \frac{n}{2j}\Big); \tag{20}$$
$$K_{ijn} = [C_{i0j0}^{n0}]^2;$$
$$\alpha_{ijn} = -\sqrt{i(i+1)j(j+1)}\,C_{i0j0}^{n0}C_{i(-1)j1}^{n0}.$$

Here, C_{i0j0}^{n0} and $C_{i(-1)j1}^{n0}$ are the Clebsch–Gordan coefficients. They are other than zero if the subscripts meet the conditions

$$|i = j| \le n \le (i+j); \quad (i+j+n) = 2g. \tag{21}$$

Therefore, only those oscillation modes whose numbers meet (21) will be excited in the second-order approximation.

(4) From the second part of (20), it follows that, if any three surface oscillation modes with numbers p, q, and k meet one of the relationships

$$\omega_p + \omega_q = \omega_k, \quad \omega_k + \omega_q = \omega_p, \tag{22}$$

then, these modes start resonantly interacting, according to the general idea of the method of many scales. In this case, one may speak of the secondary (i.e., second-order) Raman resonance.

Note that, according to (16), the frequencies ω_n of the surface eigenmodes depend on the self-charge of the drop (i.e., on parameter W). For $W_{\text{cr}} = 4$, the fundamental mode ($n = 2$) frequency vanishes, and a further increase in W makes the surface of the drop unstable against the self-charge. Therefore, the secondary resonances affect the nonlinear oscillations of the drop and merit study only if relationships (22) are valid at $W < W_{\text{cr}}$. Such a resonance was first discovered [2] in the case $\omega_6 = 2\omega_4$, and later it was shown [11, 13, 15] that the number of such resonances at $W < 4$ is large (reaches several hundreds for $p, q, k < 100$).

Let subscript n refer to the second-order modes that are excited because of nonlinear interaction and subscripts k, p, and q, to the resonantly coupled modes.

(i) Consider the case when $n \ne k, p, q$ (i.e., when mode n drops out of any resonance relationship) and the condition of eliminating secular and small-denominator terms from a solution to Eq. (20) has the simple form

$$\frac{dA_n^{(1)}(T_1)}{dt} = 0.$$

Substituting $A_n^{(1)}(T_1)$ expressed through scalar functions $a_n^{(1)}(T_1)$ and $b_n^{(1)}(T_1)$ (see (18)) into this equality and requiring that the real and imaginary parts of the resulting expression vanish, one can easily check that

$$\frac{da_n^{(1)}(T_1)}{dt} = \frac{db_n^{(1)}(T_1)}{dt} = 0.$$

These equalities imply that $a_n^{(1)}(T_1)$ and $b_n^{(1)}(T_1)$ are independent of slow time T_1 and may be taken to be equal to their constant initial values (see (19)) in the second-order approximation. First-order coefficients $M_n^{(1)}(t)$ (see expression (17)) in expansion (15) of the perturbation $\xi^{(1)}(\theta, t)$ of the equilibrium surface shape in Legendre polynomials will take the form

$$M_n^{(1)}(t) = \delta_{n,i} h_i \cos(\omega_i t); \quad i \in \Xi; \quad n \neq k, p, q, \tag{23}$$

where $\delta_{n,i}$ is the Kronecker symbol.

The amplitudes of the second-order corrections that are obtained by solving Eq. (20) will then have the form

$$\begin{aligned} M_n^{(2)}(t) &= \sum_{i \in \Xi} \sum_{j \in \Xi} h_i h_j \left\{ \lambda_{ijn}^{(+)} \sin\left(\frac{1}{2}(\omega_n + \omega_i + \omega_j)t\right) \right. \\ &\times \sin\left(\frac{1}{2}(\omega_n - \omega_i - \omega_j)t\right) + \lambda_{ijn}^{(-)} \sin\left(\frac{1}{2}(\omega_n + \omega_i - \omega_j)t\right) \\ &\left. \times \sin\left(\frac{1}{2}(\omega_n - \omega_i + \omega_j)t\right) \right\} \quad (n \geq 2;\ n \neq q, p, k); \end{aligned} \tag{24}$$

$$\lambda_{ijn}^{(\pm)} \equiv (\gamma_{ijn} \pm \omega_i \omega_j \eta_{ijn})(\omega_n^2 - (\omega_i \pm \omega_j)^2)^{-1}.$$

(ii) To analyze Eq. (20) for $n = k, p, q$, we introduce an offset parameter $\sigma \sim O(1)$ that reflects the proximity of the difference $\omega_p - \omega_q$ to frequency ω_k and can be determined from the expression

$$\omega_p - \omega_q = \omega_k(1 + \sigma k). \tag{25}$$

Note that one may relate the offset parameter to the self-charge of the drop (or to parameter W), implying that the oscillation frequency may be controlled and, in particular, detuned from the exact resonance frequency by varying the self-charge of the drop.

With (25) substituted into (20), the right of (20) contains the terms involving the following factors:

$$\begin{aligned} \exp(i(\omega_p - \omega_q)T_0) &= \exp(i(\omega_k + \varepsilon\omega_k\sigma)T_0) \\ &= \exp(i\sigma\omega_k T_1)\exp(i\omega_k T_0); \end{aligned}$$

$$\begin{aligned} \exp(i(\omega_k + \omega_q)T_0) &= \exp(i(\omega_p - \varepsilon\omega_k\sigma)T_0) \\ &= \exp(-i\sigma\omega_k T_1)\exp(i\omega_p T_0); \end{aligned}$$

$$\begin{aligned} \exp(i(\omega_p - \omega_k)T_0) &= \exp(i(\omega_q + \varepsilon\omega_k\sigma)T_0) \\ &= \exp(i\sigma\omega_k T_1)\exp(i\omega_q T_0). \end{aligned}$$

Then, for $n = k, p, q$, the conditions for eliminating secular terms from a solution to (20) can be written as

$$\begin{aligned} &-2i\omega_k \frac{dA_k^{(1)}(T_1)}{dt} + \Lambda_{pqk}^{(-)} \exp(i\sigma\omega_k T_1) A_p^{(1)}(T_1) \overline{A_q^{(1)}(T_1)} = 0; \\ &-2i\omega_p \frac{dA_p^{(1)}(T_1)}{dt} + \Lambda_{kqp}^{(+)} \exp(-i\sigma\omega_k T_1) A_k^{(1)}(T_1) \overline{A_q^{(1)}(T_1)} = 0; \\ &-2i\omega_q \frac{dA_q^{(1)}(T_1)}{dt} + \Lambda_{pkq}^{(-)} \exp(i\sigma\omega_k T_1) A_p^{(1)}(T_1) \overline{A_k^{(1)}(T_1)} = 0; \end{aligned} \tag{26}$$

$$\Lambda_{lmn}^{(\pm)} = (\gamma_{lmn} + \gamma_{mln}) \pm \omega_l \omega_m (\eta_{lmn} + \gamma_{mln}).$$

Equating the real and imaginary parts of expressions (26) to zero and introducing function

$$\beta_k^{(1)}(T_1) = \sigma\omega_k T_1 - b_k^{(1)}(T_1), \tag{27}$$

we arrive at a set of differential equations for real functions $a_k^{(1)}(T_1)$, $\beta_k^{(1)}(T_1)$, $a_p^{(1)}(T_1)$, $b_p^{(1)}(T_1)$, $a_q^{(1)}(T_1)$, and $b_q^{(1)}(T_1)$:

$$\begin{aligned} &2\omega_k \frac{da_k^{(1)}(T_1)}{dT_1} = \Lambda_{pqk}^{(-)} a_p^{(1)}(T_1) a_q^{(1)}(T_1) \sin(\varphi_{kpq}^{(1)}(T_1)); \\ &2\omega_k a_k^{(1)}(T_1) \frac{d\beta_k^{(1)}(T_1)}{dT_1} \\ &= 2\omega_k^2 a_k^{(1)}(T_1)\sigma + \Lambda_{pqk}^{(-)} a_p^{(1)}(T_1) a_q^{(1)}(T_1) \cos(\varphi_{kpq}^{(1)}(T_1)); \\ &2\omega_p \frac{da_p^{(1)}(T_1)}{dT_1} = -\Lambda_{kqp}^{(+)} a_k^{(1)}(T_1) a_q^{(1)}(T_1) \sin(\varphi_{kpq}^{(1)}(T_1)); \\ &2\omega_p a_p^{(1)}(T_1) \frac{db_p^{(1)}(T_1)}{dT_1} \\ &= -\Lambda_{kqp}^{(+)} a_k^{(1)}(T_1) a_q^{(1)}(T_1) \cos(\varphi_{kpq}^{(1)}(T_1)); \\ &2\omega_q \frac{da_q^{(1)}(T_1)}{dT_1} = \Lambda_{pkq}^{(-)} a_p^{(1)}(T_1) a_k^{(1)}(T_1) \sin(\varphi_{kpq}^{(1)}(T_1)); \end{aligned} \tag{28}$$

$$2\omega_q a_q^{(1)}(T_1)\frac{db_q^{(1)}(T_1)}{dT_1}$$
$$= -\Lambda^{(-)}_{pkq} a_p^{(1)}(T_1) a_k^{(1)}(T_1)\cos(\varphi^{(1)}_{kpq}(T_1));$$
$$\varphi^{(1)}_{kpq}(T_1) = \beta_k^{(1)}(T_1) + b_p^{(1)}(T_1) - b_q^{(1)}(T_1).$$

Relationships (19) serve as initial conditions for Eqs. (28). From the requirement that set (28) be self-consistent at $t = 0$, we find that, if one of modes k, p, and q is lacking in the spectrum Ξ of the initially excited modes (i.e., its amplitude equals zero at the zero time), its phase at $t = 0$ equals $\pi/2$ rather than being arbitrary. Eventually, the initial conditions for set (28) can be written in the compact form

$$a_j^{(1)}(0) = \delta_{i,j}h_j/2; \quad b_j^{(1)}(0) = \pm(1-\delta_{i,j})\pi/2; \qquad (29)$$
$$i \in \Xi; \quad j = k, p, q.$$

For resonantly interacting modes k, p, and q, the first-order coefficients in expansion (15) are written as (see (17))

$$M_k^{(1)}(t) = 2a_k^{(1)}(\varepsilon t)\cos((\omega_p - \omega_q)t - \beta_k^{(1)}(\varepsilon t));$$
$$M_p^{(1)}(t) = 2a_p^{(1)}(\varepsilon t)\cos(\omega_p t + b_p^{(1)}(\varepsilon t)); \qquad (30)$$
$$M_q^{(1)}(t) = 2a_q^{(1)}(\varepsilon t)\cos(\omega_q t + b_q^{(1)}(\varepsilon t)),$$

where coefficients $a_k^{(1)}(T_1)$, $\beta_k^{(1)}(T_1)$, $a_p^{(1)}(T_1)$, $b_p^{(1)}(T_1)$, $a_q^{(1)}(T_1)$, and $b_q^{(1)}(T_1)$ are the solutions to set (28) with boundary conditions (29).

It should be noted that, in the approximation used (up to the second order of smallness), three modes resonantly interact only if at least two of them are present in the spectrum Ξ of the modes excited at the zero time; that is, the amplitudes of these modes must be other than zero at $t = 0$. The third mode, even having the zero initial amplitude, appears in the first-order spectrum when its number satisfies the conditions $p + q + k$ is an even number and $|p - q| \le k \le (p + q)$ (for $p, q \in \Xi$; $k \notin \Xi$). Both conditions result from the requirement that coefficients $\Lambda^{(-)}_{pqk}$, $\Lambda^{(+)}_{kqp}$, and $\Lambda^{(-)}_{pkq}$ in Eqs. (28) vanish.

Figure 1 shows the evolution of the first-order amplitudes of the fourth, fifth, and seventh modes resonantly interacting at $W = 1.649$ when the fourth and seventh modes are responsible for the initial deformation. The calculation was made by formulas (28)–(30) for $\varepsilon = 0.3$. It is seen that the fifth mode, which is absent in the initial spectrum, results from resonant energy transfer from the seventh (highest) mode. It is also seen that the energy of the seventh mode is partially transferred to the fourth mode as well, the amplitude of which grows synchronously with the amplitude of the fifth mode. In other words, we observe energy transfer from the higher mode to the lower ones according to the concept of disintegration instability.

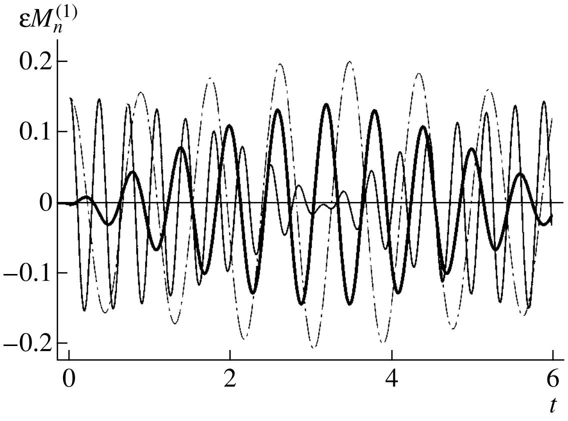


Fig. 1. Dimensionless amplitudes $M_n^{(1)}$ of the fourth, fifth, and seventh nonlinear capillary oscillation modes of the charged drop that are in exact resonance ($W = 1.649$) vs. dimensionless time. The thin, bold, and dash-and-dot lines refer, respectively, to the seventh, fifth, and fourth modes.

(iii) Consider now the case of degenerate resonance, where one of the modes interacts twice with another, i.e., where $\omega_s = 2\omega_k$.

Proceeding in the same way as above, we obtain expressions for the first-order time-dependent coefficients in expansion (15):

$$M_s^{(1)}(t) = 2a_s^{(1)}(\varepsilon t)\cos(2\omega_s t - \beta_s^{(1)}(\varepsilon t));$$
$$M_k^{(1)}(t) = 2a_k^{(1)}(\varepsilon t)\cos(2\omega_k t + b_k^{(1)}(\varepsilon t)), \qquad (31)$$

where real functions $a_s^{(1)}(\varepsilon t)$, $\beta_s^{(1)}(\varepsilon t)$, $a_k^{(1)}(\varepsilon t)$, and $\beta_k^{(1)}(\varepsilon t)$ are solutions to the set of differential equations

$$4\omega_s\frac{da_s^{(1)}(T_1)}{dT_1} = \Lambda^{(+)}_{kks}(a_k^{(1)}(T_1))^2\sin(\varphi^{(1)}_{sk}(T_1));$$

$$4\omega_s a_s^{(1)}(T_1)\frac{d\beta_s^{(1)}(T_1)}{dT_1} = 4\omega_s^2 a_s^{(1)}(T_1)\sigma$$
$$+ \Lambda^{(+)}_{kks}(a_k^{(1)}(T_1))^2\cos(\varphi^{(1)}_{sk}(T_1));$$

$$2\omega_k\frac{da_k^{(1)}(T_1)}{dT_1} = -\Lambda^{(-)}_{skk}a_s^{(1)}(T_1)a_k^{(1)}(T_1)\sin(\varphi^{(1)}_{sk}(T_1));$$

$$2\omega_k a_k^{(1)}(T_1)\frac{db_k^{(1)}(T_1)}{dT_1} \qquad (32)$$

$$= -\Lambda^{(-)}_{skk}a_s^{(1)}(T_1)a_k^{(1)}(T_1)\cos(\varphi^{(1)}_{sk}(T_1));$$
$$\varphi^{(1)}_{sk}(T_1) = \beta_s^{(1)}(T_1) + 2b_k^{(1)}(T_1);$$

$$\beta_s^{(1)}(T_1) = \sigma\omega_s T_1 - b_s^{(1)}(T_1).$$

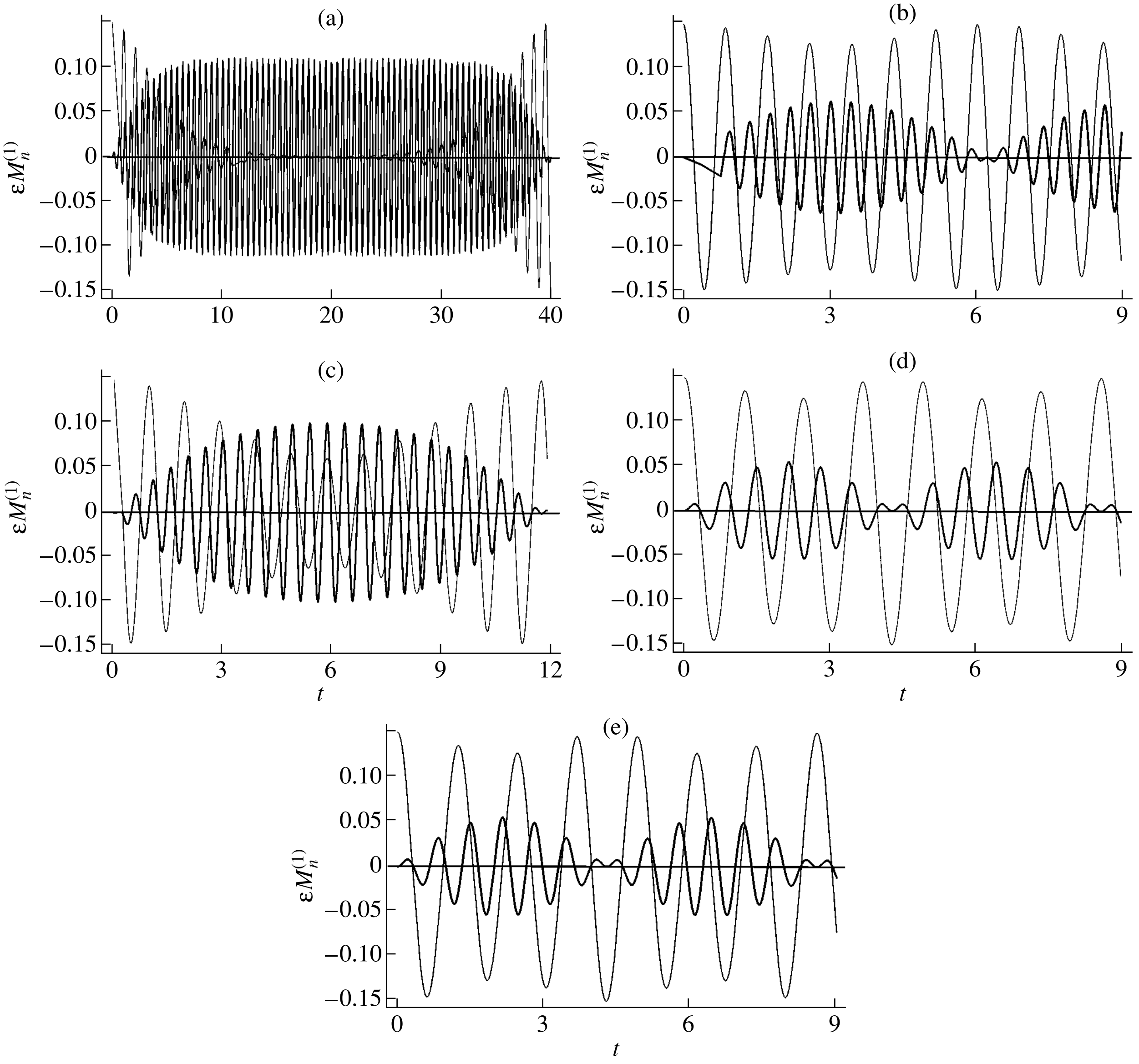


Fig. 2. Dimensionless amplitudes of the resonantly interacting fourth and sixth modes vs. dimensionless time. W = (a) 2.66667 (exact resonance), (b) 1.5, (c) 2.5, (d) 3.0, and (e) 3.9. The thin and bold lines refer, respectively, to the fourth and sixth modes.

From relationships (19), it follows that set (32) allows the following combinations of the initial conditions:

$$[s, k] \in \Xi:\ a_s^{(1)}(0) = h_s/2; \quad \beta_s^{(1)}(0) = 0;$$

$$a_k^{(1)}(0) = h_k/2; \quad b_k^{(1)}(0) = 0;$$

$$s \neq \Xi,\ k \in \Xi:\ a_s^{(1)}(0) = 0; \quad \beta_s^{(1)}(0) = \pi/2;$$

$$a_k^{(1)}(0) = h_k/2; \quad b_k^{(1)}(0) = 0.$$

If $k \notin \Xi$ and $s \in \Xi$ (i.e., if $a_k^{(1)}(0) = 0$ and $a_s^{(1)}(0) = h_s/2$), modes s and k do not resonantly interact in the approximation used, since it follows from set (32) at $t = 0$ that

$$\frac{da_s^{(1)}(0)}{dT_1} = \frac{da_k^{(1)}(0)}{dT_1} = 0,$$

that is, amplitudes $a_k^{(1)}$ and $a_s^{(1)}$ retain their initial values.

Figure 2a shows the time evolutions of the amplitudes $M_4^{(1)}(t)$ and $M_6^{(1)}(t)$ of the resonantly interacting fourth and sixth modes in the position of exact resonance, $W_r = 2.66667$, for $\varepsilon = 0.3$. Here, at the zero time, the fourth mode alone is excited, while the sixth mode is absent (its amplitude is zero). Note that, if the sixth mode alone is excited at the zero time, the resonant buildup of the fourth mode does not take place. Figures 2b–2e demonstrate the same dependences for different (other than W_r) values of parameter W, which defines offset σ.

It follows from Fig. 2 that the nonlinear interaction between the modes is of a resonance character for any $W < W_{cr} = 4$. This means that the frequency offset is small when W varies in the range considered. An increase in the absolute value of the offset parameter causes a decrease in (i) the time of resonant interaction, which depends on the time it takes for the mode amplitude to reach a maximal value; (ii) the time period over which the energy is in a resonantly growing mode; and (iii) the part of the energy that is transferred from an initially excited mode to that growing resonantly (the complete energy exchange is observed only at the exact resonance). To this must be added that the mode of initially zero amplitude acquires a first-order amplitude upon resonant buildup, although the resonant interaction itself shows up only in the second order of smallness.

(5) The conclusion that the resonance conditions depend on the self-charge of the drop only slightly can be generalized for the case of several resonance interactions proceeding simultaneously [20]. Let any, e.g., a jth mode be involved in several resonant interactions at $W < 4$ that differ in interacting modes and the values of W corresponding to exact resonances. For example, the jth mode participates in two resonance situations: j, i, k at $W_r = C_1$ and j, n, m at $W_r = C_2$, where C_1, $C_2 < 4$. Then, the excitation of the jth mode causes the modes involved in both resonance situations (ith, kth, nth, and mth) to resonantly interact with it. The amplitudes of the modes resonantly building up through the interaction with the jth mode in either combination will depend on offset parameter W in a given situation (i.e., on the difference between W and its resonant values C_1 and C_2). For example, among the first ten modes, the fourth one may take part in the following resonant interactions: at $W = 0.612$, it may resonantly interact with the sixth and eighth modes; at $W = 1.649$, with the fifth and seventh modes; at $W = 2.66667$, twice with the sixth mode (the generate resonance considered above); and at $W = 3.623$, with the third and fifth modes [16, 17]. Thus, at any $W < 4$, the fourth mode may virtually interact with all the modes listed above, with the degree of interaction (the fraction of the transferred energy) being dependent on the offset in each of the possible combinations.

Consider the situation where mode k participates in two resonance interactions (one two-mode degenerate and one three-mode nondegenerate) simultaneously. Let the offset parameters for these resonance situations be σ_1 and σ_2:

$$\omega_p - \omega_q = \omega_k(1 + \varepsilon\sigma_1); \quad 2\omega_k = \omega_s(1 + \varepsilon\sigma_2).$$

Analyzing this situation in the same way as above, we find that the first-order amplitudes for modes p, q, and k have the form of (28). For mode s, we get

$$M_s^{(1)}(t) = 2a_s^{(1)}(\varepsilon t)\cos(2(\omega_p - \omega_q)t - \beta_s^{(1)}(\varepsilon t)).$$

Functions $\beta_s^{(1)}(\varepsilon t)$ from this expression and $\beta_k^{(1)}(\varepsilon t)$ from (28) are defined as

$$\beta_k^{(1)}(T_1) = \sigma_1\omega_k T_1 - b_k^{(1)}(T_1);$$

$$\beta_s^{(1)}(T_1) = (\sigma_2\omega_s + 2\sigma_1\omega_k)T_1 - b_s^{(1)}(T_1).$$

The set of differential equations for real functions $a_p^{(1)}(\varepsilon t)$, $b_p^{(1)}(\varepsilon t)$, $a_q^{(1)}(\varepsilon t)$, $b_q^{(1)}(\varepsilon t)$, $a_s^{(1)}(\varepsilon t)$, $\beta_s^{(1)}(\varepsilon t)$, $a_k^{(1)}(\varepsilon t)$, and $\beta_k^{(1)}(\varepsilon t)$ includes the third, fourth, fifth, and sixth equations of set (28), as well as the equations

$$2\omega_k\frac{da_k^{(1)}(T_1)}{dT_1} = \Lambda_{pqk}^{(-)}a_p^{(1)}(T_1)a_q^{(1)}(T_1)\sin(\varphi_{kpq}^{(1)}(T_1))$$
$$- \Lambda_{skk}^{(-)}a_s^{(1)}(T_1)a_k^{(1)}(T_1)\sin(\varphi_{sk}^{(1)}(T_1));$$

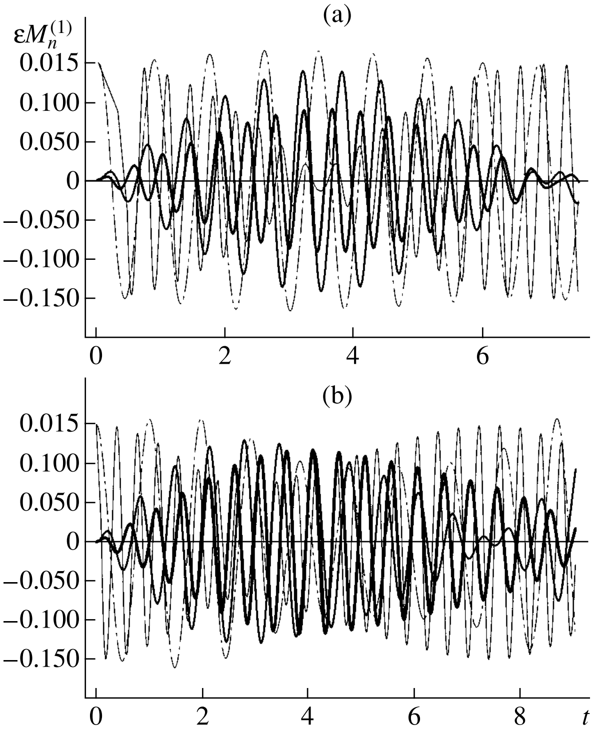


Fig. 3. Dimensionless amplitudes $M_n^{(1)}$ of the resonantly interacting fourth, fifth, sixth, and seventh modes vs. dimensionless time for W = (a) 1.649 and (b) 2.66667. The thin, heavy, bold, and dash-and-dot lines refer, respectively, to the seventh, sixth, fifth, and fourth modes.

$$2\omega_k a_k^{(1)}(T_1)\frac{d\beta_k^{(1)}(T_1)}{dT_1} = 2\omega_k^2 a_k^{(1)}(T_1)\sigma_1$$
$$+ \Lambda_{pqk}^{(-)} a_p^{(1)}(T_1) a_q^{(1)}(T_1)\cos(\varphi_{kpq}^{(1)}(T_1))$$
$$+ \Lambda_{skk}^{(-)} a_s^{(1)}(T_1) a_k^{(1)}(T_1)\cos(\varphi_{sk}^{(1)}(T_1));$$

$$4\omega_s\frac{da_s^{(1)}(T_1)}{dT_1} = \Lambda_{kks}^{(+)}(a_p^{(1)}(T_1))^2\sin(\varphi_{sk}^{(1)}(T_1)); \quad (33)$$

$$4\omega_s a_s^{(1)}(T_1)\frac{d\beta_s^{(1)}(T_1)}{dT_1} = 4\omega_s(\sigma_2\omega_s + 2\sigma_1\omega_k)a_s^{(1)}(T_1)$$
$$+ \Lambda_{kks}^{(+)}(a_k^{(1)}(T_1))^2\cos(\varphi_{sk}^{(1)}(T_1));$$
$$\varphi_{sk}^{(1)}(T_1) = \beta_s^{(1)}(T_1) - 2b_k^{(1)}(T_1).$$

The initial conditions for this set are given by (29) where $j = k, p, q, s$.

Figure 3a shows the time dependences of the resonantly interacting modes, including the resonantly growing fifth and sixth modes, for the same initial conditions as in Fig. 1. It is remembered that, in Fig. 1, the fourth, fifth, and seventh modes are under the exact resonance conditions and the initial deformation is specified by the fourth and seventh modes. The curves in Fig. 3 were calculated using set (33) complemented by the third, fourth, fifth, and sixth equations of set (28). It is seen that the energy is transferred from the seventh mode to all the lower number modes. It is interesting that, being in degenerate resonance with the fourth mode, the sixth mode builds up via taking the energy of the fourth mode [10] (see also Fig. 2). Nevertheless, as follows from Fig. 3, the fourth-mode amplitude does not decrease; on the contrary, it even slightly increases synchronously with the fifth and sixth modes. In other words, the energy transferred from the seventh mode to the fourth one not only compensates for the energy losses of the fourth mode (a part of its energy was gained by the sixth mode) but even increases its energy.

Figure 3b shows the results of the same calculations for $W = 2.66667$, i.e., when the fourth and sixth modes come into exact degenerate resonance (the initial conditions are the same as in Fig. 3a). Here, unlike Fig. 3a, both the fourth and the seventh modes give up their energy and the time dependences of the amplitudes of the resonantly growing fifth and sixth modes become asymmetric.

CONCLUSIONS

When Rayleigh parameter W, which is related to the self-charge of the drop) is subcritical ($W < 4$ for the fundamental mode), the frequency offset of the modes being excited is insufficient for nonlinearly interacting modes to be in resonance at any W irrespectively of the exact resonance values of W (W_r). In this case, the offset influences only the fraction of the energy being transferred and the transfer time.

When the self-charge of the drop is small, disintegration of a nonlinearly oscillating drop may be related to the resonant transfer of the capillary oscillation energy from higher to lower modes.

Under the three-mode resonance conditions, disintegration instability is observed only for Raman resonances. For degenerate resonances, this effect is absent.

ACKNOWLEDGMENTS

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

REFERENCES

1. J. A. Tsampopoulos and R. A. Brown, J. Fluid Mech. **127**, 519 (1983).
2. J. A. Tsampopoulos and R. A. Brown, J. Fluid Mech. **147**, 373 (1984).
3. R. Natarajan and R. A. Brown, Phys. Fluids **29**, 2788 (1986).
4. R. Natarajan and R. A. Brown, J. Fluid Mech. **183**, 95 (1987).
5. R. Natarajan and R. A. Brown, Proc. R. Soc. London, Ser. A **410**, 209 (1987).
6. Z. C. Feng and L. G. Leal, Phys. Fluids A **5**, 826 (1993).
7. Z. C. Feng and L. G. Leal, Phys. Fluids **7**, 1325 (1995).
8. Z. C. Feng and Y. H. Su, Phys. Fluids **9**, 519 (1997).
9. Z. C. Feng, J. Fluid Mech. **333**, 1 (1997).
10. S. O. Shiryaeva, Pis'ma Zh. Tekh. Fiz. **26** (22), 76 (2000) [Tech. Phys. Lett. **26**, 1016 (2000)].
11. S. O. Shiryaeva, A. I. Grigor'ev, and D. F. Belonozhko, Pis'ma Zh. Tekh. Fiz. **29** (6), 69 (2003) [Tech. Phys. Lett. **29**, 249 (2003)].
12. S. O. Shiryaeva, Zh. Tekh. Fiz. **72** (4), 15 (2002) [Tech. Phys. **47**, 389 (2002)].
13. S. O. Shiryaeva, D. F. Belonozhko, and A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **28** (22), 45 (2002) [Tech. Phys. Lett. **28**, 945 (2002)].
14. S. O. Shiryaeva, Zh. Tekh. Fiz. **73** (2), 19 (2003) [Tech. Phys. **48**, 152 (2003)].
15. S. O. Shiryaeva, A. I. Grigor'ev, and D. F. Belonozhko, Pis'ma Zh. Tekh. Fiz. **29** (6), 69 (2003) [Tech. Phys. Lett. **29**, 249 (2003)].
16. S. O. Shiryaeva, Pis'ma Zh. Tekh. Fiz. **29** (17), 28 (2003) [Tech. Phys. Lett. **29**, 714 (2003)].
17. S. O. Shiryaeva, A. N. Zharov, and A. I. Grigor'ev, Zh. Tekh. Fiz. **74** (1), 10 (2004) [Tech. Phys. **49**, 8 (2004)].
18. A. I. Grigor'ev, S. O. Shiryaeva, and M. V. Volkova, Zh. Tekh. Fiz. **73** (11), 31 (2003) [Tech. Phys. **48**, 1389 (2003)].
19. M. I. Rabinovich and D. I. Trubetskov, *Introduction to the Theory of Oscillations and Waves* (Nauka, Moscow, 1984; Kluwer, Dordrecht, 1989).
20. L. M. Brekhovskikh and V. V. Goncharov, *Introduction to the Mechanics of Continuous Media* (Nauka, Moscow, 1982) [in Russian].

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 193–196. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 53–56.
Original Russian Text Copyright © 2005 by Fateev.

SOLIDS

Low-Viscosity Flow of a Solid Dispersed Medium in the Bridgman Effect

E. G. Fateev
Institute of Applied Mechanics, Ural Division, Russian Academy of Sciences, Izhevsk, 426067 Russia
e-mail: fateev@udman.ru

Received December 1, 2003; in final form, April 20, 2004

Abstract—The quasi-fluid flow of microdispersed solids at the stage of rapid dynamic high-pressure loading under the conditions of the Bridgman effect is considered in terms of the hydrodynamic approach. Under these conditions, the effective dynamic viscosity of the rapidly flowing solid medium is estimated to be $\eta = 1$–3 Pa s, which is typical of only low-viscous fluids. The time of dynamic loading at the stage of rapid compression in the Bridgman effect is estimated as $t = (3$–$6) \times 10^{-6}$ s, which agrees well with the available experimental data. *© 2005 Pleiades Publishing, Inc.*

(1) Strong quasi-static compression of insulators and semiconductors between open Bridgman anvils (Fig. 1) in many cases causes volume fracture, which is accompanied by fast ejection of the material in the microdispersed state from the compression system [1–3]. Such explosive instabilities appear when the elastic energy of a body heavily compressed at high pressures ($P = 0.1$–10 GPa) is converted to the mechanical work of superfast volume relief once certain critical P–T parameters have been reached in the system [4–6]. Each material has its own characteristic pressure and temperature thresholds at which this type of instability occurs. Therefore, the Bridgman effect (BE) conditions [4] are well suited to studying the mechanical stability of insulators at high pressures. Also, the Bridgman effect gives the unique chance to estimate the ultimate mechanical stability of minerals, which may reside in the seismic areas of the Earth's crust or in rock-impact-prone underground workings, in laboratory conditions [4].

Two stages may be distinguished in the Bridgman effect. At the first stage, an insulator is slowly (quasi-statically) loaded with a rate $dP/dt < 0.1$ GPa/s to a certain threshold at which it fails throughout the volume, passing into the microdispersed state. The second stage starts when the microdispersed material becomes similar to a quasi-fluid and, being subjected to a high pressure, is explosively (totally or partially) ejected out of the compression system with a velocity $v_e = 0.5$–2.0 km/s [4].

The fast quasi-fluid flow of a material between Bridgman anvils during the ejection [6–15] allows one to use the Bridgman effect for effective activation-free modification of nonmetals. Unlike the other modification methods, such a mechanochemical high-pressure approach enables keeping the mean pressure and the ambient temperature constant during the process. In addition, it excludes the penetration of equipment-related chemical impurities into an agglomerate or alloy to be produced. Moreover, in the Bridgman effect conditions, one can easily control the mechanochemical processes, changing the excitation threshold of the explosion by varying the temperature [6–8] or placing the samples in electric fields with various amplitudes and frequencies [16–22].

The purpose of this work is to estimate the effective values of the viscosity, as well as the time of dynamic loading, of materials in the Bridgman effect. These estimates may be helpful in predicting the intensity of mechanochemical interactions in the reagents and gaining a better insight into the nature of the Bridgman effect.

(2) The superplasticity of materials under the critical conditions of nonuniform compression between Bridgman anvils may be associated with a substantial decrease in their viscosity when explosive breaking of atomic bonds results in local heating of grain boundaries [23]. In this case, the superplastic flow of a dispersed material may be in many ways similar to the viscous flow. Therefore, such a quasi-fluid flow can be described in terms of the hydrodynamic approach [24, 25]. However, as applied to the BE-related processes, this approximation seems to be formal. According to the available experimental data, materials subjected to dynamic loading under the BE conditions usually do not melt, while are substantially heated (by several tens or hundreds of degrees, depending on the conditions) [26, 27].

Consider a model of the quasi-fluid flow of a dispersed body at the instant of its fast compression between Bridgman anvils. The quasi-fluid film being compressed is between two surfaces rapidly approaching each other, which tend to squeeze the viscous quasi-fluid out of the gap between them (Fig. 2). If the sur-

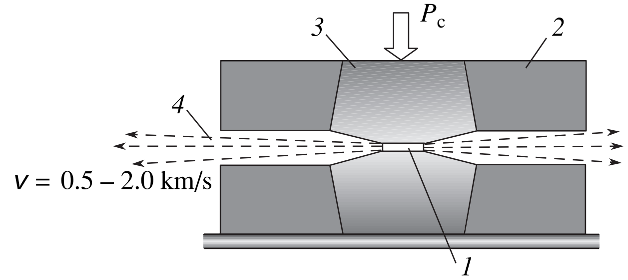


Fig. 1. Compression of sample *1* between two Bridgman anvils *2* with superhard inserts *3*, which generate the explosive effect with fast ejection *4* of the material out of the system.

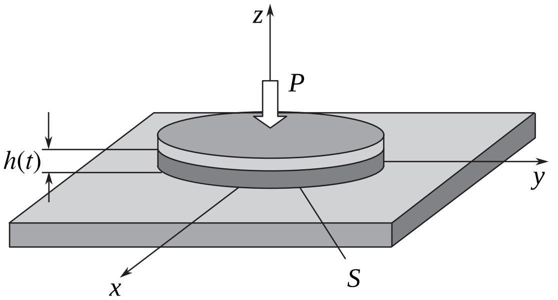


Fig. 2. Compression of a fluid film between two plates by applying a pressure *P*. The thickness *h* of the film depends on time, $h = h(t)$. *S* is the hydrodynamic contact area.

faces are flat and move parallel to each other along the normal to them and if the fluid between them is incompressible and has an effective dynamic viscosity η, the pressure distribution can be described by a set of Navier–Stokes equations with appropriate boundary conditions. In the general case, when the velocity and pressure fields are nonstationary and depend on time only through a generalized coordinate *h* (the thickness of the fluidlike film), this set of equations is reduced to the Reynolds equation. Solving it, we can find the velocity v_n of the approaching plates as a function of initial load *P* and the effective viscosity [28],

$$v_n = \frac{Ph^3}{2\eta S}. \tag{1}$$

Here, the hydrodynamic contact area is equal to the effective cross-sectional area $S = (3/4)\pi R^2$ of the anvils of radius *R*. The time it takes for the quasi-fluid film thickness to decrease from initial thickness h_0 to an intermediate thickness *h* (at a given load *P*) is found from the simple relationship

$$t = \frac{\eta S}{P}\left(\frac{1}{h^2} - \frac{1}{h_0^2}\right). \tag{2}$$

Characteristic values of the parameters in Bridgman-effect experiments

Parameter	Typical value	Refs.
h_0	$\approx 4 \times 10^{-4}$ m	[4–22]
h	$\sim 10^{-4}$ m	[4–22]
P_c	$\sim 10^9$ Pa	[4–22]
R	$\approx 2.5 \times 10^{-3}$ m	[15–22]
v_e^0	$(0.5–1.5) \times 10^3$ m/s	[4]
$v_n \approx (h_0/R)v_e^0$	$(0.8–2.4) \times 10^2$ m/s	–
η	0.8–2.6 Pa s	(6)
t	$(3–6) \times 10^{-6}$ s	(2)

Note that the viscosity may increase considerably with pressure *P* and substantially drop with increasing temperature *T*. The estimates that follow are based on the qualitative relationship

$$\eta \sim \eta_0 \exp\left(\frac{PV_0}{kT}\right), \tag{3}$$

which, in particular, can be applied for describing the viscosity of melts [29]. Here, *k* is the Boltzmann constant and V_0 is the volume of a microvoid that coincides (in order of magnitude) with the volume occupied by particles in the melt, $V_0 = (0.5–2.0) \times 10^{-29}$ m^3. This qualitative relationship is also applicable to a flow of disperse particles if we assume that the surfaces of micrometer particles are molten at the instant of ejection.

(3) Let us estimate the effective dynamic viscosity of materials rapidly flowing between the anvils at the instant the Bridgman effect is observed. At normal temperature, the ejection velocity $v_e^0 \approx (R/h)v_n$ is proportional to the velocity v_n of the anvils (see Eq. (1)). Then, we have

$$\eta_0 \sim \frac{2}{3}\frac{Ph^2}{\pi R v_e^0}. \tag{4}$$

Using the available experimental data for the parameters relevant to the conditions of Bridgman instability (see table), we obtain $\eta_0 = 0.8–2.6$ Pa s (the effect of high pressure on the dynamic viscosity is disregarded). This value is typical of a low-viscosity fluid, such as glycerin (~0.85 Pa s), in order of magnitude; is three orders of magnitude higher than the viscosity of water at its melting point (≈0.002 Pa s [30]); and is five to eight orders of magnitude lower than the viscosity of a molten oxide glass (~10^9 and ~10^6 Pa s, respectively, for SiO_2 and B_2O_3 [29]). Next, for typical values of the parameters, the time of dynamic loading (without regard to the effect of high pressure) is $t_0 = (1–4) \times 10^{-7}$ s, as follows from Eq. (2). This value is one order of magnitude lower than the estimate given in [4]. This time is likely to control the rates of mass transfer and chemical reactions. Such an assumption is also based on estimates of the relaxation times in chemical transformations observed in the Bridgman effect, ~10^{-6} s [4, 5, 9–13].

Fractionation of chemical elements under the BE conditions [15] may indicate that microparticles enriched by different elements that have entered into the quasi-fluid flow with different rates. This supposition may explain the fractionation in $GeSe_2$ disk samples [15], where the local Se concentration at the periphery of the disk that remained in the anvils after the explosive effect was 10–15% higher than the Se concentration at the center. It is worth comparing the melting points of Se, T_{Se} = 170–217°C, and Ge, T_{Ge} = 936°C. This observation means that the viscosity of Se-enriched fine particles may be much lower than that of Ge-enriched particles at the same pressures and temperatures (see Eq. (3)).

Because of low values of effective dynamic viscosity, a material placed in the BE conditions may become heterogeneous; i.e., it may so happen that the material in the quasi-fluid stratifies and some of the layers flow with a still lower viscosity. Then, it becomes clear that the reason for an extremely high rate of various chemical reactions and other interactions is the ultrahigh mobility of the elements under high pressures. It is quite possible that the reactions in the BE conditions will proceed to completion, as observed in [9–13]. Certainly, the rate of the reactions is bound to depend on the initial temperature of the reaction mixture, since the threshold pressure of the Bridgman effect is temperature dependent. To a first approximation, this dependence appears as [6, 20, 22]

$$P(T) \sim P_0(1 - AT), \tag{5}$$

where P_0 is the BE threshold at room temperature and A is a parameter depending on the structure of the material, energy parameters, and so on.

Therefore, according to Eqs. (3)–(5), the effective dynamic viscosity of the material may depend on temperature and pressure as

$$\eta \sim \frac{2P(T)h^2}{3\pi R v_e^0} \exp\frac{P(T)V_0}{kT}, \tag{6}$$

and vary as shown in Fig. 3 for the parameters listed in the table and $A \sim 10^3$ Pa/K. The flow rate of the quasi-fluid between the anvils (and, correspondingly, the ejection velocity) versus the initial temperature of the sample with regard to a high pressure in the loading system may be estimated from Eqs. (1), (5), and (6) in the form

$$v_e \sim v_e^0 \exp^{-1}\left(\frac{P(T)V_0}{kT}\right). \tag{7}$$

The run of this dependence is shown in Fig. 4. Using Eqs. (2), (5), and (6), one can also find the impact loading time with allowance for the high-pressure effect, $t = (3–6) \times 10^{-6}$ s. In order of magnitude, this value agrees well with the time found experimentally [4].

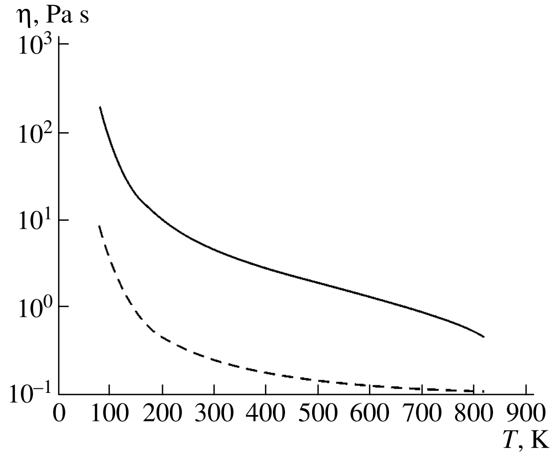


Fig. 3. Temperature dependence of the effective dynamic viscosity for the rapid flow between Bridgman anvils with (solid line) and without (dashed line) allowance for the effect of pressure on the viscosity.

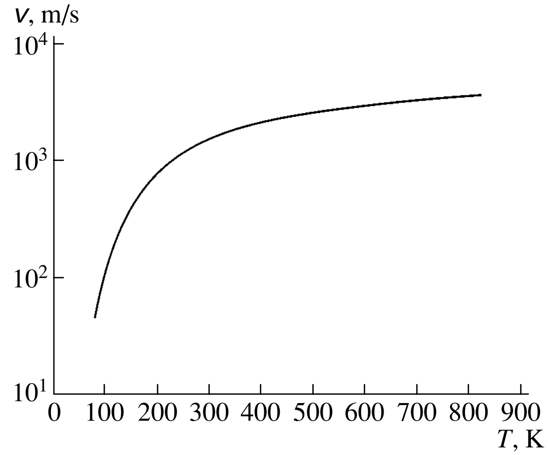


Fig. 4. Effect of the initial temperature on the flow rate of the quasi-fluid between Bridgman anvils with allowance for the high pressure effect.

(4) Thus, studying the low-viscous flow of solids in the microdispersed state under the conditions of the Bridgman effect, we managed to estimate their effective viscosity, η = 1–3 Pa s. In other words, under extremal conditions, the effective viscosity of solid dispersed media is comparable to that of low-viscous fluids, such as glycerin. Basically, the mobility of chemical elements may be extremely high under such conditions, including high pressures. A high mobility of the elements is likely to greatly intensify chemical reactions and other interactions under the conditions of the Bridgman effect, as was demonstrated in [9–15].

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 05-05-64641.

REFERENCES

1. P. V. Bridgman, Phys. Rev. **48**, 825 (1935).
2. P. V. Bridgman, Proc. Am. Acad. Arts Sci. **71**, 387 (1937).
3. P. V. Bridgman, J. Appl. Phys. **8**, 328 (1937).
4. M. A. Yaroslavskiĭ, *Rheological Explosion* (Nauka, Moscow, 1982) [in Russian].
5. N. S. Enikolopyan, A. A. Mkhitaryan, A. S. Karagezyan, *et al.*, Dokl. Akad. Nauk SSSR **292**, 887 (1987).
6. E. G. Fateev and V. P. Khan, Pis'ma Zh. Tekh. Fiz. **17** (20), 52 (1991) [Sov. Tech. Phys. Lett. **17**, 736 (1991)].
7. V. P. Khan and E. G. Fateev, Pis'ma Zh. Tekh. Fiz. **16** (8), 81 (1990) [Sov. Tech. Phys. Lett. **16**, 317 (1990)].
8. E. G. Fateev and V. P. Khan, in *Proceedings of the International Conference "Noncrystalline Semiconductors-89," Uzhgorod, 1989*, pp. 121–123.
9. N. S. Enikolopyan, V. B. Vol'eva, A. A. Khardzhyan, *et al.*, Dokl. Akad. Nauk SSSR **292**, 1165 (1987).
10. N. S. Enikolopyan, Dokl. Akad. Nauk SSSR **283**, 897 (1985).
11. N. S. Enikolopyan, A. A. Khzardzhyan, É. É. Gasparyan, *et al.*, Dokl. Akad. Nauk SSSR **294**, 1151 (1987).
12. N. S. Enikolopyan, A. A. Mkhitaryan, and A. S. Karagezyan, Dokl. Akad. Nauk SSSR **288**, 657 (1986).
13. N. S. Enikolopyan, Dokl. Akad. Nauk SSSR **302**, 630 (1988).
14. S. M. Walley, J. E. Balzer, W. G. Proud, *et al.*, Proc. R. Soc. London, Ser. A **456**, 1483 (2000).
15. E. G. Fateev and V. P. Khan, Fiz. Goreniya Vzryva, No. 2, 70 (1992).
16. E. G. Fateev, Zh. Tekh. Fiz. **66** (6), 93 (1996) [Tech. Phys. **41**, 571 (1996)].
17. E. G. Fateev, Pis'ma Zh. Tekh. Fiz. **19** (10), 48 (1993) [Tech. Phys. Lett. **19**, 313 (1993)].
18. E. G. Fateev, Pis'ma Zh. Tekh. Fiz. **20** (20), 83 (1994) [Tech. Phys. Lett. **20**, 847 (1994)].
19. E. G. Fateev, Dokl. Akad. Nauk **354**, 252 (1997).
20. E. G. Fateev, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 876 (1997) [JETP Lett. **65**, 919 (1997)].
21. E. G. Fateev, Zh. Tekh. Fiz. **71** (6), 37 (2001) [Tech. Phys. **46**, 682 (2001)].
22. E. G. Fateev, Zh. Tekh. Fiz. **73** (4), 43 (2003) [Tech. Phys. **48**, 421 (2003)].
23. Yu. I. Krasnoshchekov, L. K. Kuznetsov, V. N. Perevezentsev, *et al.*, Dokl. Akad. Nauk SSSR **312**, 872 (1990) [Sov. Phys. Dokl. **35**, 573 (1990)].
24. M. A. Lavrent'ev, Usp. Mat. Nauk **12** (6), 41 (1957).
25. A. S. Balankin, Pis'ma Zh. Tekh. Fiz. **14**, 1231 (1988) [Sov. Tech. Phys. Lett. **14**, 534 (1988)].
26. K. B. Abramova, A. B. Pakhomov, B. P. Peregud, *et al.*, Zh. Tekh. Fiz. **58**, 817 (1988) [Sov. Phys. Tech. Phys. **33**, 497 (1988)].
27. A. I. Pekhovich and V. M. Zhidkikh, *Thermal Analysis of Solids* (Énergiya, Leningrad, 1976) [in Russian].
28. M. A. Galakhov, P. B. Gusyatnikov, and A. P. Novikov, *Mathematical Models of Contact Hydrodynamics* (Nauka, Moscow, 1985) [in Russian].
29. D. S. Sanditov and G. M. Bartenev, *Physical Properties of Disordered Systems* (Nauka, Moscow, 1982) [in Russian].
30. D. Eĭzenberg and V. Kautsman, *Structure and Properties of Water* (Gidrometeoizdat, Leningrad, 1975) [in Russian].

Translated by K. Shakhlevich

Technical Physics, Vol. 50, No. 2, 2005, pp. 197–201. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 57–61.
Original Russian Text Copyright © 2005 by Khishchenko, Zhernokletov, Lomonosov, Sutulov.

SOLIDS

Dynamic Compressibility, Release Adiabats, and the Equation of State of Stilbene at High Energy Densities

K. V. Khishchenko*, M. V. Zhernokletov, I. V. Lomonosov*, and Yu. N. Sutulov****

** Institute of High Energy Densities, Associated Institute for High Temperatures,*
Russian Academy of Sciences, Moscow, 125412 Russia
e-mail: konst@ihed.ras.ru
*** All-Russia Research Institute of Experimental Physics (VNIIEF),*
Russian Federal Nuclear Center, Sarov, Nizhni Novgorod Oblast, 607200 Russia

Received April 6, 2004; in final form, July 6, 2004

Abstract—The shock adiabat of porous stilbene (1,2-diphenylethylene) up to the pressure $P = 41$ GPa and the dynamic compressibility of this material in reflected shock waves up to 77 GPa are studied experimentally. The run of the expansion isentropes of stilbene down to 0.1 GPa is determined. The experimental findings are used to construct a semi-empiric equation of state of stilbene for a wide range of high-energy states. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

The equation of state of a material exposed to pulsed high-energy actions is of great interest for high energy density physics [1–4]. The behavior of materials in compression shock waves and adiabatic expansion waves provides valuable information for checking the adequacy of theoretical models or describing the thermodynamics of extreme states on a semi-empirical basis over a wide range of the phase diagram [1, 3, 4].

In this work, we report data for the compressibility of porous stilbene $C_6H_5CH{=}CHC_6H_5$ (1,2-diphenylethylene) in shock waves of intensity up to $P = 41$ GPa, as well as systematize the states of the material upon repeat dynamic loading up to $P = 77$ GPa and in waves of adiabatic unloading down to $P = 0.1$ GPa. The earlier [5] and newly found data are generalized in the form of a semi-empiric equation of state.

EXPERIMENTAL METHODS AND RESULTS

Test porous samples (cylindrical pellets) were prepared by pressing of flakes. The density of stilbene crystals under normal conditions is $\rho_0 = 1.16$ g/cm^3. In our experiments, it varied from $\rho_{00} = 1.13$ to 0.85 g/cm^3. Shock waves were generated through copper, aluminum, or iron screens by steel striking rods, which were accelerated to 5–6 km/s by the detonation products of condensed explosives. With the shock wave amplitude in the screen fixed, determination of wave velocity D (accurate to ≈1.5%) in the sample makes it possible to find mass velocity U and pressure P from a known dynamic adiabat of the screen by the reflection method [4]. The measurements were made by the contact electrical method, and the signals from the sensors were recorded on a fast-response oscilloscope. In the experiments, the diameter and thickness of the samples were 12 and 3 mm, respectively. Table 1 lists the dynamic compressibility values for stilbene with $\rho_{00} = 0.85$ g/cm^3 that were averaged over six to eight independent measurements, and Fig. 1 shows the associated data points in D–U coordinates.

In the experiments with reflected shock waves and adiabatic expansion waves, the initial density of the samples was $\rho_{00} = 1.13$ g/cm^3. The velocity D_* of the

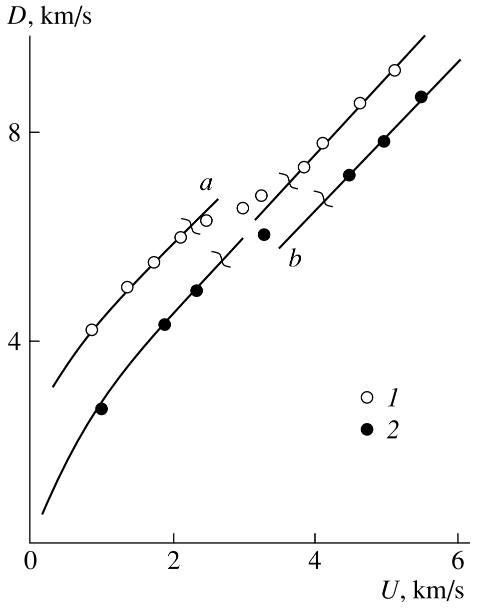


Fig. 1. Shock adiabats for the stilbene samples with the initial porosity $m_p = \rho_0/\rho_{00} =$ (*a*) 1.027 and (*b*) 1.365. Wavy segments show the onset of physicochemical transformation. Data points are taken from (*1*) [5] and (*2*) this work.

1063-7842/05/5002-0197$26.00 © 2005 Pleiades Publishing, Inc.

Table 1. Experimental data for the shock compressibility of Stilbene

U, km/s	D, km/s	P, GPa	ρ, g/cm^3
0.98	2.64	2.199	1.3518
1.85	4.29	6.746	1.4945
2.34	4.97	9.885	1.6063
3.24	6.06	16.689	1.8266
4.44	7.24	27.324	2.1979
4.95	7.89	33.197	2.2811
5.47	8.79	40.869	2.2505

Note: ρ_{00} = 0.85 g/cm^3, m_p = 1.365.

shock waves was measured in obstacles with different dynamic impedances that were placed behind the samples. From these measurements, the parameters of the initial states were found. Data points on the repeat shock adiabats were obtained by reflection from obstacles (aluminum, Teflon) for which the dynamic adiabats are known and that are more rigid than the test material. The low pressure (high rarefaction) range was studied with soft dynamic obstacles (polyethylene, polystyrene foamed plastic of different density, as well as argon and air under atmospheric pressure) for which the shock adiabats are well known (see, e.g., [6]). In most of the experiments, the 2-mm-thick test samples of 38-mm diameter were covered not only by 3-mm-thick pellets of 12-mm diameter made of the obstacle materials but also by pellets made of the test material itself in order to keep track of the initial state parameters. The parameters of the initial and final states in the condensed obstacles were measured using the reflection method with the shock wave velocity recorded by contact electrical sensors. In the case of argon, the optical radiation emitted from the shock wave front was recorded by optical methods. When the air obstacle was used for relieving, the velocity U of the free surface of the sample upon passing the first shock wave through it was measured by the electrical contact sensors. In this case, the wave front velocity D_* in the air obstacle was measured, and pressure P behind the front was found from the known shock adiabat for air. The experimental techniques and approaches used in this paper were described at greater length in [6], where isentropic expansion of shock-compressed materials was considered. The experimental data for repeat shock compression and adiabatic expansion of stilbene are summarized in Table 2, and the data points, each being the average of six to eight independent measurements, are applied on the P–U diagram (Fig. 2).

MODEL OF THE EQUATION OF STATE

On the pressure–volume–energy surface, the early [5] and newly obtained shock-wave data for stilbene show the run of only two shock adiabats (for the samples with ρ_{00} = 1.13 and 0.85 g/cm^3). Our data for repeat shock compression and isentropic expansion shed new light upon the stilbene behavior at high energy densities. Similar experiments performed earlier on metals [3, 6–9], Plexiglas and Teflon [10], polystyrene and Fenilon [11, 12], and carbon [13, 14] confirm the validity of the thermodynamic description of a new (previously uncovered) range of densities and pressures. However, a bare handful of experimental data and some theoretical uncertainty regarding the behavior of organic substances at high energy densities impose substantial restrictions on the form of the semi-empiric potential and necessitate using a simplified model of the equation of state with few adjustable parameters.

In generalized form, the caloric model of the equation of state [11, 15], in terms of which the thermody-

Table 2. Experimental data for repeat shock compression and adiabatic expansion of stilbene

Material of obstacle	Adiabats R_1 and S_1			Adiabats R_2 and S_2			Adiabats R_3 and S_3		
	D_*, km/s	U, km/s	P, GPa	D_*, km/s	U, km/s	P, GPa	D_*, km/s	U, km/s	P, GPa
Aluminum	8.97	2.68	65.147	9.22	2.87	71.710	9.42	3.01	76.840
Teflon	6.82	2.91	43.463	7.41	3.26	52.903	8.22	3.75	67.507
Stilbene	7.394	3.83	32.0	7.878	4.10	36.5	8.598	4.58	44.5
Polyethylene	–	–	–	8.95	4.12	33.924	–	–	–
Foamed plastic 0.74	7.11	4.31	22.677	–	–	–	8.26	5.12	31.295
Foamed plastic 0.39	6.91	5.04	13.582	7.23	5.29	14.916	8.02	5.88	1.840
Foamed plastic 0.16	–	–	–	8.02	6.29	8.071	–	–	–
Argon	–	–	–	9.04	8.05	0.130	–	–	–
Air	–	–	–	8.64	8.23	0.092	–	–	–

Note: Figures by the foamed plastics are densities in g/cm^3. ρ_{00} = 1.13 g/cm^3, m_p = 1.027.

namic properties of stilbene are described, is given by

$$P(V, E) = P_{\mathrm{c}}(V) + \frac{\Gamma(V, E)}{V}[E - E_{\mathrm{c}}(V)], \qquad (1)$$

where $E_{\mathrm{c}}(V)$ and $P_{\mathrm{c}}(V) = -dE_{\mathrm{c}}/dV$ are the elastic components of energy and pressure at $T = 0$ K and $\Gamma(V, E)$ is the coefficient that takes into account the contribution of thermal components to the equation of state.

The volume dependence of the elastic compression energy is represented in the potential [7–9]

$$E_{\mathrm{c}}(V) = \frac{B_{0\mathrm{c}}V_{0\mathrm{c}}}{m-n}(\sigma_{\mathrm{c}}^{m}/m - \sigma_{\mathrm{c}}^{n}/n) + E_{\mathrm{d}}, \qquad (2)$$

where $\sigma_{\mathrm{c}} = V_{0\mathrm{c}}/V$, $V_{0\mathrm{c}}$ is the specific volume at $P = 0$ and $T = 0$ K, and $B_{0\mathrm{c}}$ is the value of bulk modulus $B_{\mathrm{c}} = -VdP_{\mathrm{c}}/dV$ at $\sigma_{\mathrm{c}} = 1$.

Quantity E_{d}, which has the meaning of the characteristic energy of destruction, is found from the normalization condition $E_{\mathrm{c}}(V_{0\mathrm{c}}) = 0$, which yields $E_{\mathrm{d}} = B_{0\mathrm{c}}V_{0\mathrm{c}}/mn$. The derivative of the bulk modulus with respect to pressure, $B'_{0\mathrm{c}} = dB_{\mathrm{c}}/dP_{\mathrm{c}}$, at $\sigma_{\mathrm{c}} = 1$ provides a relationship between coefficients m and n in the form $n = B'_{0\mathrm{c}} - m - 2$. Parameters $V_{0\mathrm{c}}$, $B_{0\mathrm{c}}$, and $B'_{0\mathrm{c}}$ are found by iteration so as to fit (under normal conditions) the tabulated value of specific volume $V = V_0$, as well as adiabatic sound velocity $C_S = C_{S0}$ and the derivative of isentropic bulk modulus $B_S = -V(\partial P/\partial V)_S = C_S^2/V$ with respect to pressure ($B'_S = (\partial B_S/\partial P)_S = B'_{S0}$), which are determined from dynamic measurements. Exponent m in formula (2) is taken such that the resulting equation of state provides the best fit to the experimental data obtained in forward and backward shock waves.

Similarly to the caloric model [7, 8], the dependence of coefficient Γ on volume and energy is stated as

$$\Gamma(V, E) = \gamma_i + \frac{\gamma_{\mathrm{c}}(V) - \gamma_i}{1 + \sigma^{-2/3}[E - E_{\mathrm{c}}(V)]/E_{\mathrm{a}}}, \qquad (3)$$

here, $\sigma = V_0/V$, function $\gamma_{\mathrm{c}}(V)$ refers to the range of low thermal energies, and γ_i characterizes a heavily heated condensed substance (high energy range). Energy of anharmonicity E_{a} specifies the thermal energy of transition from one limiting case to the other and is found from data of dynamic high-pressure experiments.

The volume dependence of the elastic component γ_{c} of Γ is given by [11, 15]

$$\gamma_{\mathrm{c}}(V) = 2/3 + (\gamma_{0\mathrm{c}} - 2/3)\frac{\sigma_n^2 + \ln^2\sigma_m}{\sigma_n^2 + \ln^2(\sigma/\sigma_m)}, \qquad (4)$$

where

$$\gamma_{0\mathrm{c}} = \gamma_i + (\gamma_0 - \gamma_i)\left[1 + \frac{E_0 - E_{\mathrm{c}}(V_0)}{E_{\mathrm{a}}}\right]^2,$$

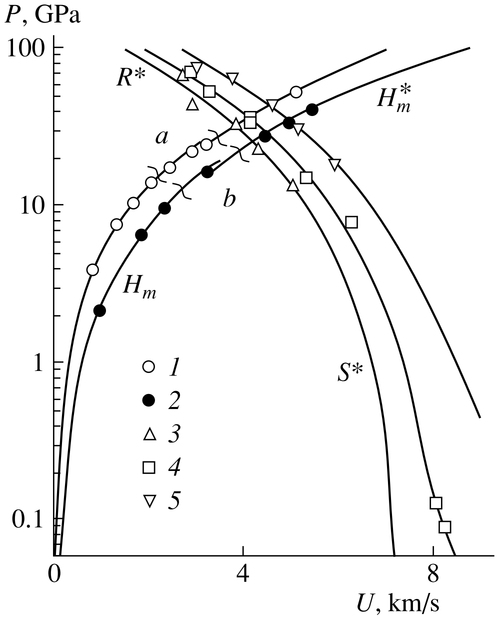


Fig. 2. Experimental (data points) and calculated (lines) shock adiabats (H_m), repeat compression curves (R), and expansion isentropes (S) of stilbene. The curves marked by the asterisk refer to the transformation products. (*3*) Adiabats R_1 and S_1; (*4*), R_2 and S_2; and (*5*) R_3 and S_3. For (*1*) and (*2*), see Fig. 1.

E_0 is the specific internal energy, and γ_0 is Grüneisen coefficient $\gamma = V(\partial P/\partial E)_V$ under normal conditions.

It is easy to check that dependence (4) provides the fulfillment of the condition $\gamma(V_0, E_0) = \gamma_0$, as well as yields the asymptotics $\gamma_{\mathrm{c}} = 2/3$ in the limiting cases of low and high compressions. Adjustable parameters σ_n and σ_m in expression (4) are taken so as to provide an optimal fit to experimentally found dynamic compressibilities and release isentropes of the porous samples.

RESULTS OF CALCULATION

The experimental data for stilbene obtained earlier [5] and in this work suggest that the material undergoes a physicochemical transformation at the front of intense shock waves. In the shock adiabat for the solid samples, this transformation starts at $P \approx 15$ GPa and results in a considerable decrease in the density (by ≈20%) and compressibility. Such transformations, which are triggered by intense dynamic loading and cause a drastic consolidation of the material, are typical of the class of aromatic substances [11, 12, 16–22]. They are routinely explained by bond breaking in the initial compound with the formation of a slightly compressible mixture of diamond-like carbon and various low-molecular components [23]. In this work, the thermodynamic properties of stilbene are studied in terms of caloric model (1)–(4) separately for the states before and after the transformation. The density and bulk modulus of the transformation products under normal con-

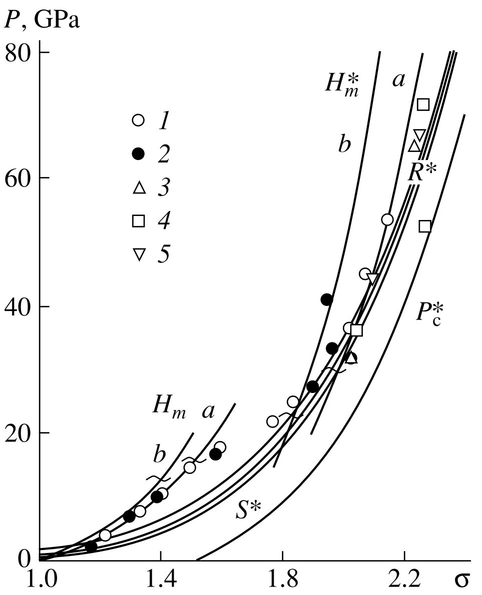


Fig. 3. Phase diagram of stilbene at high densities and pressures. For designations, see Fig. 2. P_c is the elastic compression curve at $T = 0$ K.

ditions, which are necessary for constructing the equation of state, were found from the shock-wave experiments performed at pressures above the pressure range of the transformation. The coefficients of the equations of state that optimally generalize the available thermodynamic data for stilbene and for the products of its transformation are given in Table 3.

As follows from Fig. 1, the equations of state derived in this work fit well the shock compressibility of the solid and porous stilbene samples throughout the range of the kinematic parameters. When calculating the parameters of the shock adiabats corresponding to the material transformed, we took into account the effective initial porosity of stilbene, which may influence the density of the transformation products. It should be noted that the combined analysis of the calculated adiabats and experimental data made it possible to find the parameters of state of the transformation products under normal conditions with a high degree of reliability.

Table 3. Coefficients in the equation of state for stilbene

V_0	V_{0c}	B_{0c}	m	n	σ_m	σ_n	γ_{0c}	γ_i	E_a
0.862	0.857	12.136	2.3	0.9	0.8	1	0.7	0.5	30
0.57*	0.568	28.7	3.95	1.05	0.575	1	0.75	0.5	70

Note: Coefficients in the lower row, which is marked by asterisk, refer to the products of physicochemical transformation of stilbene under shock wave loading. The coefficients are given in the basic units of measure: $P = 1$ GPa, $V = 1$ cm^3/g, and $E = 1$ kJ/g

Figure 2 shows the calculated shock adiabats for stilbene, as well as the curves of repeat loading and the expansion isentropes for the transformation products. Comparing the calculation results with the experimental data, we can argue that the equations of state constructed reliably determine the material parameters at pressures higher and lower relative to the shock adiabat. The same is true for the whole range of densities and pressures that was used in the compression experiments, as follows from Fig. 3, which compares the calculated adiabats with the data points obtained in the forward and backward (reflected) shock waves.

The P–σ diagram depicted in Fig. 3 demonstrates the domain of high-energy states occurring in the reflected compression and unloading (release) waves versus the parameters of the shock adiabats for the samples of different initial porosity. The extent of the domain is considerable both toward cold curve P_c (repeat adiabats R) and toward the rarefied gaseous phase (isentropes S). It should be noted that the initial states for the release isentropes are the dissociated states, which arise under intense dynamic loading. These isentropes persist up to rarefied gaseous states at a pressure $P \approx 0.1$ GPa and density $\rho \approx 0.3\rho_0$. In this parameter domain, any noticeable steps in the thermodynamic functions or any hydrodynamic anomalies that might be related to phase transitions in the products of stilbene transformation in the release wave were not observed (Fig. 2).

To conclude, our experimental data cover the previously unstudied range of densities and pressures. The constructed equations of state for stilbene and the products of its transformation fit well the available data gained in high-energy experiments. Therefore, these wide-range equations will be helpful in numerically simulating the effect of intense energy fluxes on a material.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research, grant. no. 03-02-16687.

The authors also thank the Russian Science Support Foundation.

REFERENCES

1. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Nauka, Moscow, 1966; Academic, New York, 1967).
2. *The Physics of High Energy Densities*, Ed. by P. Caldirola and H. Knoepfel (Academic, New York, 1971; Mir, Moscow, 1974).
3. A. V. Bushman, V. E. Fortov, G. I. Kanel', and A. L. Ni, *Intense Dynamic Loading of Condensed Matter* (Taylor & Francis, Washington, 1993).

4. *Shock Waves and Extremal States of Substance*, Ed. by V. E. Fortov, L. V. Al'tshuler, R. F. Trunin, and A. I. Funtikov (Nauka, Moscow, 2000) [in Russian].
5. K. V. Khishchenko, M. V. Zhernokletov, I. V. Lomonosov, and Yu. N. Sutulov, Khim. Fiz. **17**, 29 (1998).
6. L. V. Al'tshuler, A. V. Bushman, M. V. Zhernokletov, *et al.*, Zh. Éksp. Teor. Fiz. **78**, 741 (1980) [Sov. Phys. JETP **51**, 373 (1980)].
7. A. V. Bushman and I. V. Lomonosov, in *Investigation of Materials Properties under Extremal Conditions*, Ed. by V. E. Fortov and E. A. Kuz'menkov (IVTAN, Moscow, 1990), pp. 34–40 [in Russian].
8. A. V. Bushman, V. E. Fortov, and I. V. Lomonosov, *High Pressure Equations of State: Theory and Applications*, Ed. by S. Eliezer and R. A. Ricci (North Holland, Amsterdam, 1991), pp. 249–262.
9. A. V. Bushman, I. V. Lomonosov, and V. E. Fortov, *Equations of State of Metals at High Energy Density* (Inst. Khim. Fiz., Chernogolovka, 1992) [in Russian].
10. A. V. Bushman, M. V. Zhernokletov, I. V. Lomonosov, *et al.*, Dokl. Akad. Nauk **329**, 581 (1993) [Phys. Dokl. **38**, 165 (1993)].
11. A. V. Bushman, M. V. Zhernokletov, I. V. Lomonosov, *et al.*, Zh. Éksp. Teor. Fiz. **109**, 1662 (1996) [JETP **82**, 895 (1996)].
12. K. V. Khishchenko, I. V. Lomonosov, V. E. Fortov, *et al.*, in *Proceedings of the 5th International Conference "Zababakhin Scientific Talks," Snezhinsk, 1998* (RFYaTs-VNIITF, Snezhinsk, 1999), pp. 714–720.
13. M. N. Pavlovskiĭ, K. V. Khishchenko, M. V. Zhernokletov, *et al.*, in *Physics of Matter under Extreme Conditions-2001*, Ed. by V. Fortov, V. P. Efremov, V. Khishchenko, *et al*. (Inst. Khim. Probl. Khim. Fiz. RAN, Chernogolovka, 2001), pp. 96–97 [in Russian].
14. K. V. Khishchenko, V. E. Fortov, I. V. Lomonosov, *et al.*, in *Shock Compression of Condensed Matter-2001*, Ed. by M. D. Furnish, N. N. Thad-hani, and Y. Horie (AIP, New York, 2002), pp. 759–762.
15. I. V. Lomonosov, V. E. Fortov, and K. V. Khishchenko, Khim. Fiz. **14**, 47 (1995).
16. R. D. Dick, J. Chem. Phys. **52**, 6021 (1970).
17. R. H. Warnes, J. Chem. Phys. **53**, 1088 (1970).
18. R. F. Trunin, M. V. Zhernokletov, N. F. Kuznetsov, and Yu. N. Sutulov, Khim. Fiz. **8**, 539 (1989).
19. R. F. Trunin, M. V. Zhernokletov, V. V. Dorokhin, and N. V. Sychevskaya, Khim. Fiz. **11**, 557 (1992).
20. A. V. Bushman, V. P. Efremov, V. E. Fortov, *et al.*, in *Shock Compression of Condensed Matter-1991*, Ed. by S. C. Schmidt, R. D. Dick, J. W. Forbes, and D. G. Tasker (North Holland, Amsterdam, 1992), pp. 79–82.
21. A. V. Bushman, M. V. Zhernokletov, I. V. Lomonosov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **58**, 640 (1993) [JETP Lett. **58**, 620 (1993)].
22. K. K. Krupnikov and V. P. Krupnikova, in *Shock Waves Marseille III*, Ed. by R. Brun and L. Z. Dumitrescu (Springer-Verlag, Berlin, 1995), pp. 301–306.
23. F. H. Ree, J. Chem. Phys. **70**, 974 (1979).

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 202–206. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 62–66.
Original Russian Text Copyright © 2005 by Markov.

Effect of the Carrier Concentration Distribution on the Thermoelectric Branch Efficiency

O. I. Markov
Orel State University, Orel, 302015 Russia
e-mail: Markov@e-mail.ru

Received April 20, 2004

Abstract—The one-dimensional boundary-value problem of determining the stationary temperature field of a thermoelectric branch is numerically solved for the case of a maximal temperature difference. The Fermi energy distribution over the thermoelectric branch is calculated in terms of the quantum statistics of carriers. Homogeneous and inhomogeneous thermoelectric branches are considered. © *2005 Pleiades Publishing, Inc.*

The low efficiency of thermoelectric coolers holds back their production. Therefore, improvement of the thermoelectric figure of merit is among the most important problems of semiconductor materials science. The values of this parameter that have been reached up to now are still very far from the theoretical limits [1]. Moreover, even today's figure of merit of thermoelectric materials is not used in full measure. Application of thermoelectric materials in cooling devices depends on their operating temperature interval. Accordingly, the material of choice is that offering the highest figure of merit under given temperature conditions. Therefore, there arises the need to optimize the properties of thermoelectric materials using various optimization approaches. In this work, we consider optimization of the thermoelectric material properties in the operating temperature range of a thermoelement.

As is known, the basic quality criterion of a thermoelectric material is its figure of merit [2]

$$Z = \frac{\alpha^2\sigma}{\chi}, \tag{1}$$

where α is the differential thermal emf, σ is the electric conductivity, and χ is the thermal conductivity.

Parameter Z as a function of temperature and carrier concentration is usually calculated under the assumption that either the lattice or electron component of the thermal conductivity can be ignored. In the latter case, expressions for the kinetic effects are analytically easy in the case of nondegenerate carriers and the thermoelectric figure of merit may be calculated [2]. However, these calculations are not rigorous; rather, they are estimates of the parameter. In addition, it is good to bear in mind that the most important thermoelectric materials feature weak degeneracy in the range where the figure of merit reaches a maximum [3].

Let us find the optimum chemical potential under which Z reaches a maximum value. For this purpose, we will calculate the thermoelectric figure of merit for a one-band semiconductor for the case of moderately degenerate carriers obeying the quadratic dispersion law. From liquid-nitrogen temperature on, carrier scattering by lattice vibrations dominates even in heavily doped semiconductors. Therefore, we will consider carrier scattering only by acoustic phonons. The charge carrier mobility, lattice thermal conductivity, and the scattering parameter are assumed to be independent of the carrier concentration.

Using the one-band approximants given in [4], we calculate the differential thermal emf,

$$\alpha = -\frac{k}{e}\left(\frac{F_2(\eta)}{F_1(\eta)} - \eta\right), \tag{2}$$

charge carrier mobility,

$$u = \frac{e\tau_0(T)}{m^*}\frac{F_1(\eta)}{F_{3/2}(\eta)}, \tag{3}$$

electric conductivity,

$$\sigma = enu, \tag{4}$$

thermal conductivity,

$$\chi = \chi_{\mathrm{ph}} + L\sigma T, \tag{5}$$

and Lorentz number,

$$L = \left(\frac{k}{e}\right)^2\left[\frac{F_3(\eta)}{F_1(\eta)} - \left(\frac{F_2(\eta)}{F_1(\eta)}\right)^2\right]. \tag{6}$$

Here,

$$F_i(\eta) = \int_0^\infty\left(-\frac{\partial f_0}{\partial x}\right)x^i dx$$

are the Fermi integrals [4], where $f_0 = [1 + \exp(x - \eta)]^{-1}$ is the equilibrium distribution function. The effective mass of carriers was taken to be equal to $0.5m_0$.

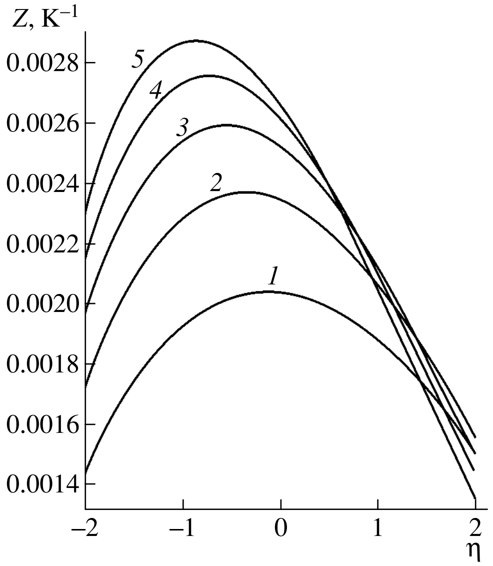


Fig. 1. Thermoelectric figure of merit vs. the reduced Fermi level at (*1*) 100, (*2*) 150, (*3*) 200, (*4*) 250, and (*5*) 300 K.

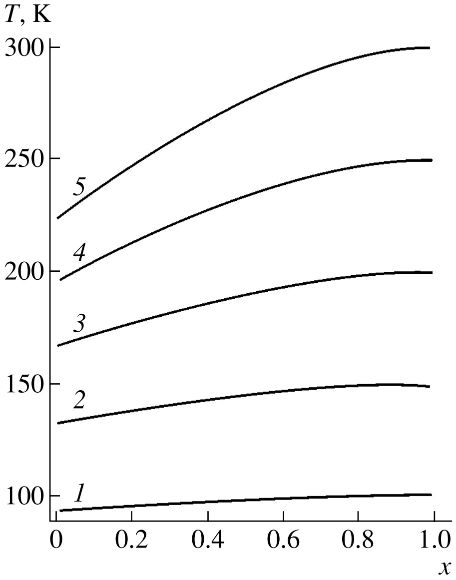


Fig. 2. Temperature distribution in the unloaded thermoelectric branch at hot end temperatures of (*1*) 100, (*2*) 150, (*3*) 200, (*4*) 250, and (*5*) 300 K.

Let the lattice (phonon) component of the thermal conductivity depend on temperature in the form

$$\chi_{\mathrm{ph}} = \mathrm{const}/T. \tag{7}$$

The absolute values of the carrier mobility and lattice thermal conductivity were taken to be close to those in the best thermoelectric materials, which have a room-temperature figure of merit near 3×10^{-3} K^{-1}. For scattering by acoustic phonons, the dependence of the carrier mobility on temperature and effective mass has the form [4]

$$u \sim m^{*-5/2} T^{-3/2}. \tag{8}$$

The results of numerical calculation of the thermoelectric figure of merit versus the reduced Fermi energy at temperatures from 100 to 300 K are presented in Fig. 1. As the temperature rises, the reduced Fermi level decreases, reaching –0.87 at room temperature. The value of the reduced Fermi level thus obtained is usually taken to be optimal.

However, one should always bear in mind that the thermoelectric figure of merit can be used only if the kinetic coefficients are temperature independent [2]; otherwise, it cannot be viewed as a reliable criterion of thermoelectric efficiency. Therefore, to calculate the thermoelectric branch efficiency, we apply another approach. Under the steady-state conditions with allowance for the Thomson effect, the temperature field of a one-dimensional adiabatically isolated homogeneous thermoelectric branch is described by the stationary equation of heat conduction

$$\frac{d}{dx}\left(\chi\frac{dT}{dx}\right) + \frac{y^2}{\sigma} - yT\frac{d\alpha}{dT}\frac{dT}{dx} = 0 \tag{9}$$

with the boundary conditions

$$\chi\frac{dT}{dx}\bigg|_{x=0} = \alpha y T|_{x=0}, \quad T|_{x=1} = T_1, \tag{10}$$

where $y = Jl/S$, l is the length of the branch, S is the cross-sectional area of the branch, and J is the current in the branch. Coefficients α, σ, and χ are calculated by formulas (2), (4), and (5), respectively.

Since the boundary-value problem given by (9) and (10) is nonlinear, it was solved numerically. Simultaneously, numerical optimization in terms of current and carrier concentration was performed. The position of the reduced Fermi level was found from the expression for the carrier concentration [4]

$$n = \frac{8\pi}{3h^3}(2m^*kT)^{3/2} F_{3/2}(\eta). \tag{11}$$

The reduced Fermi level varied between –4 and 2.5.

The temperature field and reduced Fermi level that were calculated for a thermoelectric branch operating in the conditions of maximal temperature difference are presented in Figs. 2 and 3, respectively. It is of interest to compare both approaches. From Figs. 1 and 3, it follows that the optimum values of the reduced Fermi energy, which determine the positions of the maxima of Z (Fig. 1), even fall beyond the range of variability of the reduced Fermi level along the thermoelectric branch (Fig. 3), although generally the reduced Fermi level increases with decreasing temperature. Thus, one

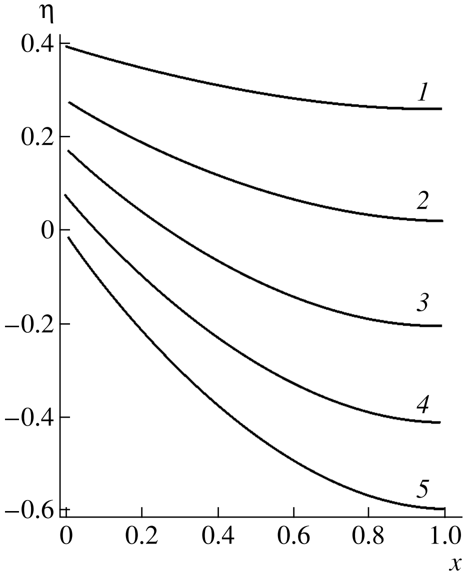


Fig. 3. Distribution of the reduced Fermi energy along the unloaded branch at different hot-end temperatures. (*1*–*5*) The same as in Fig. 2.

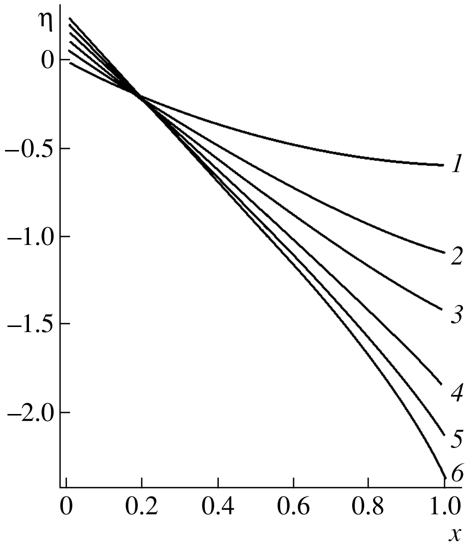


Fig. 4. Distribution of the reduced Fermi energy along the thermoelectric branch at T_1 = 300 K for the linear distribution of the carrier concentration. n_0/n_1 = (*1*) 1, (*2*) 1.5, (*3*) 2, (*4*) 3, (*5*) 4, and (*6*) 5.

can conclude that the performance of a thermoelectric branch intended for operation in a given temperature interval can be optimized by optimization of the boundary-value problem. Figure 2 shows the temperature distributions along the thermoelectric branch for the optimum values of the reduced current under the maximal temperature difference conditions. In this regime, the temperature reaches a maximum at the hot end of the branch. Therefore, the heat flux at the hottest point is absent. A vague mirror similarity between the temperature curves in Fig. 2 and the dependences of the reduced Fermi level in Fig. 3 stands out. Such behavior of the reduced Fermi level is apparently explained by the fact that the charge carrier concentration in the branch is fixed in the framework of the one-band model.

Actually, the kinetic coefficients in thermoelectric materials are temperature-dependent; therefore, in order that a thermoelectric branch be optimized in the entire temperature range, it must be inhomogeneous. The thermoelectric efficiency may be raised by using thermoelectric branches that are inhomogeneous along their length [5]. It has been established that the thermal efficiency is improved if the electric conductivity increases and the thermopower decreases from the hot to cold end. Such a conclusion was drawn from the study of graded-property thermoelements as a limiting case of composed thermoelement [6]. The electric conductivity distribution along the length that improves the branch efficiency was also found. It turned out to be linear.

The calculation of the thermoelectric branch efficiency by solving the boundary-value problem was demonstrated by Ivanova and Rivkin [7]. When solving this problem in the variational statement using the maximum principle of Pontryagin, they had to linearize the initial conditions. It was also assumed that the thermopower, thermal conductivity, and electric conductivity depend on temperature only slightly. It was found that the optimal charge concentration distribution along the branch is also linear.

The author of this paper solved this problem in a somewhat different formulation with allowance for a temperature dependence of the kinetic coefficients and the linear distribution of the charge carrier concentration [8, 9]. The regimes of maximal temperature difference and maximal cooling capacity were considered. Eventually, it was found that the linear distribution of the carrier concentration increases both the temperature difference across and the cooling capacity of a thermoelectric branch. The weak point in the studies cited is that they use classical statistics, while thermoelectric materials feature a weak degeneracy of carriers. Therefore, quantum statistics should be invoked to solve this problem. The stationary temperature field of a one-dimensional adiabatically isolated inhomogeneous thermoelectric branch is described by the stationary equation of heat conduction (with allowance for the Thomson effect and distributed Peltier effect)

$$\frac{d}{dx}\left(\chi\frac{dT}{dx}\right)+\frac{y^2}{\sigma}-yT\left(\frac{d\alpha}{dT}\frac{dT}{dx}+\frac{d\alpha}{dx}\right) = 0 \qquad (12)$$

with the same boundary conditions as (10), where $y = Jl/S$ and coefficients α, σ, and χ are calculated by formulas (2), (4), and (5), respectively.

The boundary-value problem given by (10) and (12) was solved numerically, and the solutions obtained were numerically optimized in terms of current. Since

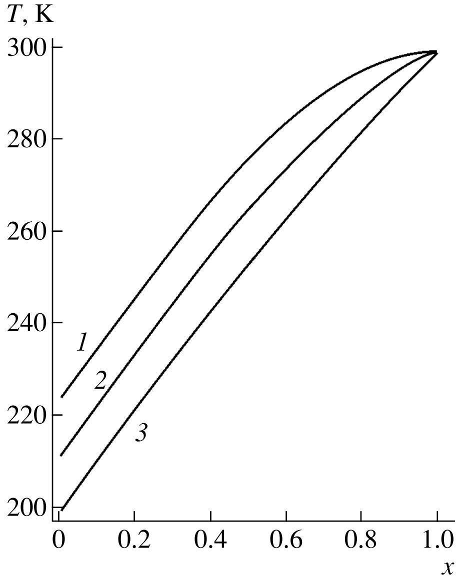


Fig. 5. Temperature distribution along the thermoelectric branch at $T_1 = 300$ K for the linear distribution of the carrier concentration. n_0/n_1 = (*1*) 1, (*2*) 2, and (*3*) 5.

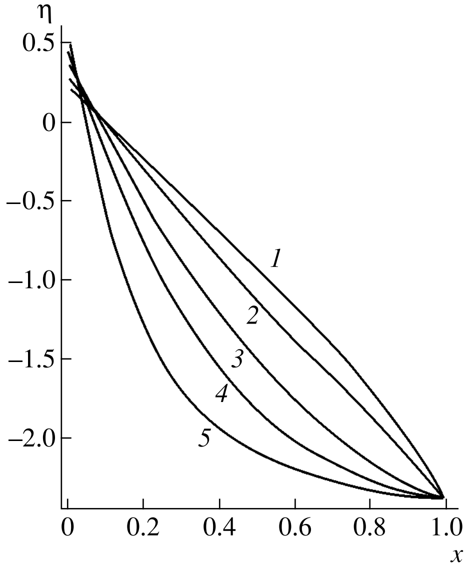


Fig. 6. Distribution of the reduced Fermi energy along the thermoelectric branch at $T_1 = 300$ K, $n_0/n_1 = 5$, and the carrier concentration distributions obeying (*1*) the linear law and exponential law with a = (*2*) 1, (*3*) 3, (*4*) 5, and (*5*) 10.

the optimal distribution of the carrier concentration for the nonlinear problem is impossible to find, we will consider particular distribution functions of charge carriers along the branch. It was argued [6, 7] that the optimal distribution is linear; therefore, the linear law of the carrier concentration distribution,

$$n = n_0(1 - gx), \tag{13}$$

will be considered first. Here, $g = 1 - 1/k$, where $k = n_0/n_1$ is the ratio of the concentrations at the cold and hot ends. Let this ratio vary in the interval $1 \le k \le 5$.

For an inhomogeneous branch, the distribution of the reduced Fermi level depends on k (Fig. 4) and so does the temperature difference, which increases with k but nonlinearly: as the concentration drop increases, the temperature difference tends to saturation. As is seen from Fig. 5 (curves *2*, *3*), the temperature is no longer maximal at the hot end of the branch and the heat flux from the hot end rises, since the temperature gradient at this point is other than zero. To avoid this, the region of heat absorption in the distributed Peltier effect should be shifted toward the cold end of the branch. Accordingly, the carrier concentration distribution must be changed. Specifically, the carrier concentration near the hot end of the branch must be kept unchanged and the region where the concentration varies sharply must be shifted to the cold end. This can be achieved if the carrier concentration distribution obeys the exponential law,

$$n = n_0(b\exp(-ax) + c), \tag{14}$$

where n_0 is the carrier concentration at the cold end of the branch, $b = (k - 1)/k(1 - \exp(-a))$, and $c = (1 - k \times \exp(-a)/k(1 - \exp(-a))$.

The distribution of the reduced Fermi energy along an exponentially inhomogeneous branch is demonstrated in Fig. 6. The concentration gradient lowers the temperature at the cold end, because a lower tempera-

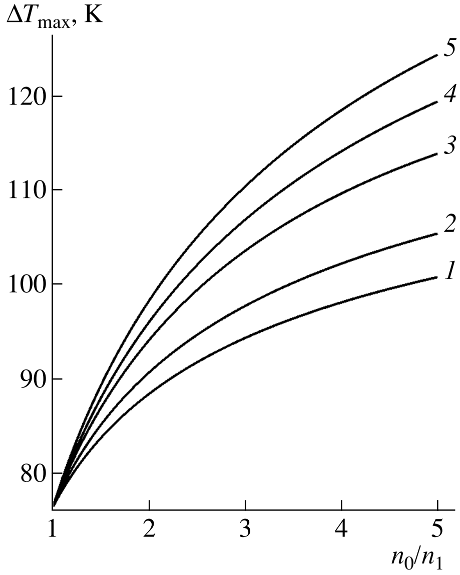


Fig. 7. Maximal temperature difference across the branch vs. the ratio n_0/n_1 of the carrier concentrations at the hot and cold ends at $T_1 = 300$ K and the carrier concentration distributions obeying (*1*) the linear law and exponential law with a = (*2*) 1, (*3*) 3, (*4*) 5, and (*5*) 10.

ture is needed in this case to compensate for the Joule heat in this region. Figure 7 plots the temperature difference against the concentration drop for various concentration distributions along the branch. Curve *1* describes the linear distribution; the other, exponential distributions with different constants *a* in the exponent. The concentration drop being the same, the maximal temperature difference increases with constant *a* in the exponent.

To conclude, we showed that the thermoelectric figure of merit as an adjustable parameter in optimization of the charge concentration can be used only at the (initial) stage of thermoelectric material selection. The final stage of optimizing the carrier concentration in the operating temperature interval of a thermoelement implies the solution of an optimal boundary-value problem even for a homogeneous branch. The linear distribution of the charge carrier concentration along the branch is not optimal: a thermoelectric branch with exponential distribution of the carrier concentration offers a higher efficiency.

REFERENCES

1. E. S. Rittner, J. Appl. Phys. **30**, 702 (1959).
2. A. F. Ioffe, *Semiconductor Thermoelements and Thermoelectric Cooling* (USSR Academy of Sciences Press, Moscow, 1960; Infosearch, London, 1957).
3. B. M. Gol’tsman, V. A. Kudinov, and I. A. Smirnov, *Bi_2Te_3-Based Thermoelectric Semiconductors* (Nauka, Moscow, 1972) [in Russian].
4. B. M. Askerov, *Electron Transport Phenomena in Semiconductors* (Nauka, Leningrad, 1985; World Sci., Singapore, 1993).
5. A. H. Boerijk, J. Appl. Phys. **32**, 1584 (1961).
6. V. A. Semenyuk, Teplofiz. Teplotekh., No. 35, 80 (1978).
7. K. F. Ivanova and F. S. Rivkin, Zh. Tekh. Fiz. **52**, 1406 (1982) [Sov. Phys. Tech. Phys. **27**, 853 (1982)].
8. O. I. Markov, Inzh.-Fiz. Zh. **76** (6), 185 (2003).
9. O. I. Markov, Zh. Tekh. Fiz. **74** (2), 138 (2004) [Tech. Phys. **49**, 277 (2004)].

Translated by Yu. Vishnyakov

Technical Physics, Vol. 50, No. 2, 2005, pp. 207–212. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 67–72.
Original Russian Text Copyright © 2005 by Emets.

Effective Permittivity of Three-Phase Composite Materials with an Anisotropic Structure

Yu. P. Emets
Institute of Electrodynamics, National Academy of Sciences of Ukraine,
Kiev, 03680 Ukraine
e-mail: emets@irpen.kiev.ua

Received April 22, 2004

Abstract—The mean electrical characteristics of an insulating matrix composite that contains unidirectional fibers of two sorts are calculated analytically. The fibers differ in physical properties and in cross-sectional area. They form alternating layers and make the material anisotropic. The calculations are based on finding an exact solution to the problem of interaction between cylindrical bodies in an external electric field. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

This paper concerns the electrical properties of a composite insulator that represents a matrix reinforced by long cylindrical fibers of two sorts having different permittivities and radii. The fibers are parallel to each other and form a cubic lattice (a square array in the cross section). Each sort of the fibers is arranged into alternating layers. Although individually the components of the composite are isotropic, the composite as a whole is anisotropic. The electrical performance of such composites is of interest in studying a number of natural objects and for developing composites to be applied in electrical engineering, mechanics, and thermal physics [1–4].

The calculation of the effective permittivity tensor is carried out by the standard procedure: first, local electric fields in the system are calculated and then they are spatially averaged. It suffices to calculate the field in one cell of the periodic structure, since the structure is regular. A cell is chosen such that its sides coincide with equipotentials and field lines. An exact solution to the problem of interaction between two parallel cylinders in an external magnetic field is of great importance in calculating the local electric field [5], which can be found in an approximation of any order. However, as the accuracy of calculation rises, the analytical expressions for the effective parameters become awkward and much more difficult to visualize. To shed light on the behavior of the mean characteristics of the structure, it suffices to consider the case of low inclusion concentrations, which simplifies the mathematics and still yields comprehensible results.

LOCAL AND MEAN ELECTRIC FIELDS

Let us consider an infinite dielectric medium of permittivity ε_1 with a doubly periodic arrangement of unidirectional cylindrical fibers of permittivities ε_2 and ε_3 and radii r_1 and r_2. The cross-sectional plane of the fibers is subdivided into square cells with period h. The fiber axes pass through the vertices of the cells. Either sort of the inclusions is arranged into rows that periodically alternate along the x axis (Fig. 1). The concentrations of the inclusions in the material, s_1 and s_2, are given by the expressions

$$s_1 = \pi r_1^2/2h^2, \quad s_2 = \pi r_2^2/2h^2. \tag{1}$$

A uniform external electric field is directed normally to the fiber axes.

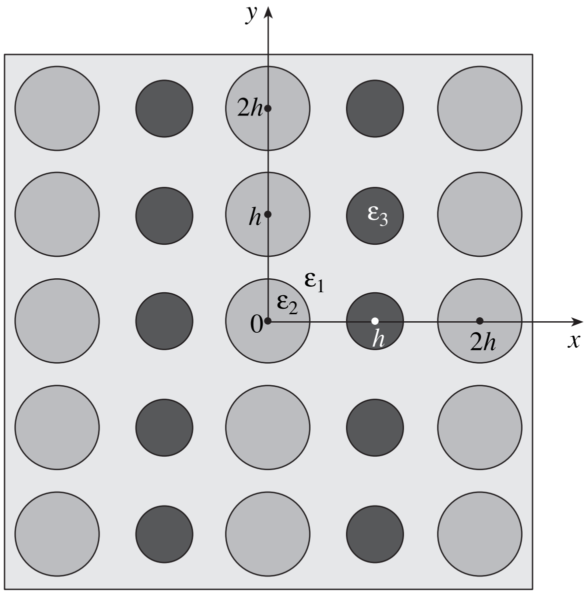


Fig. 1. Fragment of an insulating composite with cubic packing of long circular fibers (square array of the fibers in the cross section).

Under these conditions, the electric field in the material is two-dimensional and can be calculated in any approximation. The calculation scheme is as follows. The electric field inside and near an inclusion is determined as a result of its interaction with each of the other inclusions in the system. The paired interactions are then summed using an exact solution to the model problem of interaction between two parallel cylindrical bodies with arbitrary permittivities and radii that are subjected to a uniform electric field [5]. The interplay of the inclusions is mathematically expressed through dipole–dipole interaction. These are image dipoles located inside the circles that bound the cylindrical inclusions on a plane. The number of the dipoles is infinitely large, but their moments indefinitely decrease with increasing order. If the concentration of the inclusions is small, as assumed in this work, the calculations may be restricted to the one-dipole approximation. In this case, the interaction between only the first dipoles is taken into account. These dipoles are placed at the centers of the circles and have the highest moment; hence, they are the major contributors to the interplay of the inclusions in the system. To refine the results of the calculations, the second, third, etc., dipoles should be included. Their moments and coordinates are defined by specified parameters of the system and are found from simple relationships.

Following this calculation scheme, we write expressions for the electric field in a square cell. To be definite, we place the origin of the rectangular coordinate system on the axis of an inclusion with permittivity ε_2. If only the first image dipoles are taken into account, the electric field is given by the following expressions:

near the inclusion,

$$E_1(z) = E_0 - \bar{E}_0\left\{\Delta_{12}r_1^2 z^{-2} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}[\Delta_{12}r_1^2(z-a_{mn})^{-2} + \Delta_{13}r_2^2(z-b_{mn})^{-2}]\right\}; \quad (2)$$

inside the inclusion,

$$E_2(z) = (1+\Delta_{12})\left\{E_0 - \bar{E}_0\left\{\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}[\Delta_{12}r_1^2(z-a_{mn})^{-2} + \Delta_{13}r_2^2(z-b_{mn})^{-2}]\right\}\right\}. \quad (3)$$

Here,

$$E_\nu(z) = E_{x\nu} - iE_{y\nu} \quad (z = x+iy); \quad (\nu = 1, 2) \quad (4)$$

is the complex electric field intensity, $E_0 = E_{0x} - iE_{0y}$ is the external uniform electric field intensity (the bar above E_0 means complex conjugation), and

$$\Delta_{1\nu} = \frac{\varepsilon_1 - \varepsilon_\nu}{\varepsilon_1 + \varepsilon_\nu} \quad (-1 \le \Delta_{1\nu} \le 1); \quad \nu = 2, 3 \quad (5)$$

is the parameter describing the relative permittivity of each sort of the inclusions.

It follows from relationship (5) that

$$\Delta_{1\nu} = -\Delta_{\nu 1}. \quad (6)$$

In Eqs. (2) and (3), a_{mn} are the coordinates of the dipoles located at the centers of the inclusions with permittivity ε_2 and radius r_1 (except for the dipole of the inclusion in which the electric field is determined; i.e., at the origin) and b_{mn} are the coordinates of the dipoles located at the centers of the inclusions with permittivity ε_3 and radius r_2, At the x axis, we have, according to Fig. 1,

$$a_{mn} = \pm 2mh, \quad b_{mn} = \pm h(2m-1);$$

at the y axis,

$$a_{mn} = \pm 2nh;$$

outside the axes,

$$a_{mn} = h(\pm 2m \pm in), \quad b_{mn} = h[\pm(2m-1) \pm in],$$

where $m, n = 1, 2, \ldots$.

Below, when calculating the effective parameters of the system, we will need an expression for the electric field inside and near an inclusion with permittivity ε_3. In this case, one may use Eqs. (2) and (3), making the substitution

$$E_2(z) \longrightarrow E_3(z), \quad \Delta_{12} \longleftrightarrow \Delta_{13}, \quad r_1 \longleftrightarrow r_2. \quad (7)$$

Such a substitution shifts the origin to the center of an inclusion with dielectric permittivity ε_3.

Equations (2) and (3) include the interaction of the inclusions in a first approximation. If the concentration of the inclusions is so small that their interplay is insignificant, the double sums in Eqs. (2) and (3) can be neglected. As a result, we obtain

$$E_1(z) = E_0 - \bar{E}_0\Delta_{12}r_1^2 z^{-2}, \quad E_2(z) = E_0(1+\Delta_{12}). \quad (8)$$

In this case, the electric field is uniform in all the inclusions. Expressions (8) (and similar formulas for an inclusion with permittivity ε_3) define the electric field in the composite in a zeroth approximation.

To find the effective parameters of the system, it is necessary to spatially average the local electric fields in a plane normal to the fiber axes,

$$\langle \mathbf{D} \rangle = \hat{\varepsilon}_{\text{eff}} \langle \mathbf{E} \rangle. \quad (9)$$

Here, $\hat{\varepsilon}_{\text{eff}}$ is a symmetric effective permittivity tensor that has two components, $\varepsilon_{\text{eff},xx}$ and $\varepsilon_{\text{eff},yy}$.

Owing to the regular structure of the composite, it suffices to average the electric field in one cell that covers inclusions of both sorts. The cell of choice for the calculations is a rectangle like that shown in Fig. 2. If the external field is aligned with the *x* axis ($E_0 = E_{0x}$) in this case, segments *OV* and *UW* will be equipotentials and segments *OU* and *VW*, field lines. If the external field is aligned with the *y* axis ($E_0 = iE_{0y}$), segments *OU* and *VW* are equipotential lines, while *OV* and *UW* are field lines. Thus, depending on the external field direction, the potential remains unchanged at one boundaries of cell *OUWV* and the tangential electric field on the others. In this case, the calculation of the mean field simplifies and reduces to taking of contour integrals.

Let the external electric field be directed along the *x* axis in the system; that is, $E_0 = E_{0x}$. Then, the mean values of the field can be expressed by

$$\langle D_x\rangle = \frac{2}{h}\left[\varepsilon_2\int_0^{r_1}E_{2x}(y)dy + \varepsilon_1\int_{r_1}^{h/2}E_{1x}(y)dy\right],$$

$$\langle E_x\rangle = \frac{1}{h}\left[\int_0^{r_1}E_{2x}(x)dx + \int_{r_1}^{h/2}E_{1x}(x)dx + \int_{h/2}^{h-r_2}E_{1x}(x)dx + \int_{h-r_2}^{h}E_{3x}(x)dx\right]. \qquad (10)$$

Note that, in (10), the values of field $E_{1x}(x)$ differ over the intervals $[r_1, h/2]$ and $[h/2, h - r_2]$ in the same way as was mentioned when substitutions (7) were introduced. The calculations yield

$$\langle D_x\rangle = \varepsilon_1[1 - 2(2 - A_1)\Delta_{12}r_1^2 + 2B_1\Delta_{13}r_2^2 - 2\Delta_{12}^2r_1^2\Psi_1(r_1) - 2\Delta_{12}\Delta_{13}r_2^2\phi_1(r_1)]. \qquad (11)$$

Here, A_1 and B_1 are the constants that are numerically found by the formulas

$$A_1 = -4\left\{\sum_{m=1}^{\infty}\left[\frac{1}{1+16m^2} + \frac{1}{1-4m^2}\right] + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{1+2n}{(1+2n)^2+16m^2} + \frac{1-2n}{(1-2n)^2+16m^2}\right]\right\} = 0.893051, \qquad (12)$$

$$B_1 = -4\left\{\sum_{m=1}^{\infty}\left[\frac{1}{1+4(2m-1)^2}\right] + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{1+2n}{(1+2n)^2+4(2m-1)^2} + \frac{1-2n}{(1-2n)^2+4(2m-1)^2}\right]\right\} = -1.106949.$$

Functions $\Psi_1(r_1)$ and $\phi(r_1)$ are given by

$$\Psi_1(r_1) = -2\left\{\sum_{m=1}^{\infty}\left[\frac{r_1}{r_1^2+4m^2} + \frac{r_1}{r_1^2-m^2}\right] + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{r_1+n}{(r_1+n)^2+4m^2} + \frac{r_1-n}{(r_1-n)^2+4m^2}\right]\right\},$$

$$\phi_1(r_1) = -2\left\{\sum_{m=1}^{\infty}\left[\frac{r_1}{r_1^2+(2m-1)^2}\right] + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{r_1+n}{(r_1+n)^2+(2m-1)^2} + \frac{r_1-n}{(r_1-n)^2+(2m-1)^2}\right]\right\}. \qquad (13)$$

Hereafter, formulas are written in relative values

$$r_{1,2*} = \frac{r_{1,2}}{h}, \quad |E_{1,2,3*}| = \frac{|E_{1,2,3}|}{|E_0|}, \quad \varepsilon_{1,2,3*} = \frac{\varepsilon_{1,2,3}}{\varepsilon_0},$$

where ε_0 is the permittivity of vacuum (the asterisks will be omitted in subsequent expressions for brevity).

It is worth noting that formula (11) is asymmetric in equivalent parameters Δ_{12}, Δ_{13} and r_1, r_2. This is because the expression for $\langle D_x\rangle$ depends on the way this value is calculated. In (11), the flux of vector **D** on segment *OV* in the calculation cell (Fig. 2) is specified by the integrals in (10). Performing the same (equivalent) calculations on segment *UW*, we come to

$$\langle D_x\rangle = \varepsilon_1[1 - 2(2 - A_1)\Delta_{13}r_2^2 + 2B_1\Delta_{12}r_1^2 - 2\Delta_{13}^2r_2^2\Psi_1(r_2) - 2\Delta_{12}\Delta_{13}r_1^2\phi_1(r_2)]. \qquad (14)$$

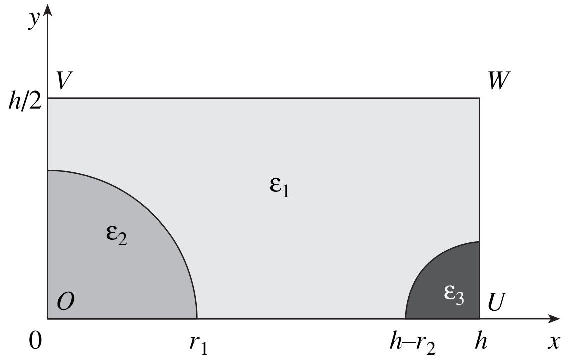


Fig. 2. Cell used to calculate the effective parameters.

Although expressions (11) and (14) differ in form, they, in fact, define the same value $\langle D_x \rangle$ (this can be checked by direct calculations) and pass into each other when the parameters are changed as $\Delta_{12} \longleftrightarrow \Delta_{13}$ and $r_1 \longleftrightarrow r_2$.

Let us determine $\langle D_x \rangle$ as an arithmetic mean,

$$\langle D_x \rangle = \frac{1}{2}[\langle D_x \rangle_{OV} + \langle D_x \rangle_{UW}], \tag{15}$$

where subscripts *OV* and *VW* indicate the ranges of integration.

Eventually, we get the expression

$$\begin{aligned} \langle D_x \rangle &= \varepsilon_1[1 - (2 - A_1 - B_1)(\Delta_{12} r_1^2 + \Delta_{13} r_2^2) \\ &- \Delta_{12}\Delta_{13}\Phi_1(r_1, r_2) - \Delta_{12}^2 r_1^2 \Psi_1(r_1) - \Delta_{13}^2 r_2^2 \Psi_1(r_2)], \end{aligned} \tag{16}$$

where

$$\Phi_1(r_1, r_2) = r_2^2 \phi_1(r_1) + r_1^2 \phi_1(r_2). \tag{17}$$

As is seen, Eq, (16) is symmetric with respect to parameters Δ_{12} and Δ_{13}, r_1 and r_2 and will be used in further calculations in such a form.

Calculations of the electric field following the scheme described above lead to the following expression:

$$\begin{aligned} \langle E_x \rangle &= 1 + (2 + A_2 + B_2)(\Delta_{12} r_1^2 + \Delta_{13} r_2^2) \\ &+ \Delta_{12}\Delta_{13}\Phi_2(r_1, r_2) + \Delta_{12}^2 r_1^2 \Psi_2(r_1) + \Delta_{13}^2 r_2^2 \Psi_2(r_2). \end{aligned} \tag{18}$$

Here, constants A_2 and B_2 are given by

$$\begin{aligned} A_2 &= 4\left\{\sum_{m=1}^{\infty} \frac{1}{1 - 4(2m-1)^2} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{1 - 2(2m-1)}{[1 - 2(2m-1)]^2 + 4n^2} + \frac{1 + 2(2m-1)}{[1 - 2(2m-1)]^2 + 4n^2}\right]\right\} = -1.390452, \\ B_2 &= 4\left\{\sum_{m=1}^{\infty} \frac{1}{1 - 16m^2} + \sum_{n=1}^{\infty} \frac{1}{1 + 4n^2} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{1 - 4m}{(1 - 4m)^2 + 4n^2} + \frac{1 + 4m}{(1 + 4m)^2 + 4n^2}\right]\right\} = 0.318397. \end{aligned} \tag{19}$$

Functions $\Phi_2(r_1, r_2)$ and $\Psi_2(r_1, r_2)$ in Eq. (18) have the form

$$\Phi_2(r_1, r_2) = r_1^2 \phi_2(r_2) + r_2^2 \phi_2(r_1),$$

$$\Psi_2(r_{1,2}) = 2\left\{\sum_{m=1}^{\infty} \frac{r_{1,2}}{r_{1,2}^2 - 4m^2} + \sum_{n=1}^{\infty} \frac{r_{1,2}}{r_{1,2}^2 + n^2} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{r_{1,2} - 2m}{(r_{1,2} - 2m)^2 + n^2} + \frac{r_{1,2} + 2m}{(r_{1,2} + 2m)^2 + n^2}\right]\right\}, \tag{20}$$

where

$$\phi_2(r_{1,2}) = 2\left\{\sum_{m=1}^{\infty} \frac{r_{1,2}}{r_{1,2}^2 - (2m-1)^2} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\frac{r_{1,2} - 2m + 1}{(r_{1,2} - 2m + 1)^2 + n^2} + \frac{r_{1,2} + 2m - 1}{(r_{1,2} + 2m - 1)^2 + n^2}\right]\right\}. \tag{21}$$

Expressions (16) and (18) allow for determination of the component $\varepsilon_{\mathrm{eff},\,xx}$ of the effective permittivity tensor. The second component of the tensor, $\varepsilon_{\mathrm{eff},\,yy}$, is found by similar calculations. One should consider the same cell *OUVW* as before, but the external electric field is now directed along the *y* axis: $E_0 = iE_{0y}$.

The mean values of the electric field, $\langle D_y \rangle$ and $\langle E_y \rangle$, are determined from the integral relations

$$\begin{aligned} \langle D_y \rangle &= \frac{1}{h}\left[\varepsilon_2 \int_0^{r_1} E_{2y}(x)dx + \varepsilon_1 \int_{r_1}^{h/2} E_{1y}(x)dx + \varepsilon_1 \int_{h/2}^{h - r_2} E_{1y}(x)dx + \varepsilon_3 \int_{h - r_2}^{h} E_{3y}(x)dx\right], \\ \langle E_y \rangle &= \frac{2}{h}\left[\int_0^{r_1} E_{2y}(y)dy + \int_{r_1}^{h/2} E_{1y}(y)dy\right]. \end{aligned} \tag{22}$$

Following the above computational scheme, we obtain

$$\begin{aligned} \langle D_y \rangle &= \varepsilon_1[1 - (2 + A_2 + B_2)(\Delta_{12} r_1^2 + \Delta_{13} r_2^2) \\ &+ \Delta_{12}\Delta_{13}\Phi_2(r_1, r_2) + \Delta_{12}^2 r_1^2 \Psi_2(r_1) + \Delta_{13}^2 r_2^2 \Psi_2(r_2)], \\ \langle E_y \rangle &= 1 + (2 - A_1 - B_1)(\Delta_{12} r_1^2 + \Delta_{13} r_2^2) \\ &- \Delta_{12}\Delta_{13}\Phi_1(r_1, r_2) - \Delta_{12}^2 r_1^2 \Psi_1(r_1) - \Delta_{13}^2 r_2^2 \Psi_1(r_2). \end{aligned} \tag{23}$$

It is easy to see that the mean values of the field satisfy the symmetry transformations [6]

$$\begin{aligned} \langle D_x \rangle(\Delta_{12}\Delta_{13}) &= \varepsilon_1 \langle E_y \rangle(\Delta_{21}, \Delta_{31}), \\ \langle D_y \rangle(\Delta_{12}\Delta_{13}) &= \varepsilon_1 \langle E_x \rangle(\Delta_{21}, \Delta_{31}). \end{aligned} \tag{24}$$

These relationships make it possible to check the validity of calculation.

EFFECTIVE PERMITTIVITY TENSOR

In the cross section of the fiber axes, the averaged properties of the material are described by the symmetric effective permittivity tensor

$$\hat{\varepsilon}_{\text{eff}} = \begin{vmatrix} \varepsilon_{\text{eff},xx} & 0 \\ 0 & \varepsilon_{\text{eff},yy} \end{vmatrix}. \qquad (25)$$

Using formulas (16), (18), and (23), we can find the components of tensor $\hat{\varepsilon}_{\text{eff}}$ that appear in Eq. (9),

$$\begin{aligned} \varepsilon_{\text{eff},xx} &= \varepsilon_1 \frac{1 - 2\alpha(\Delta_{12}s_1 + \Delta_{13}s_2) - \Delta_{12}\Delta_{13}\Phi_1(r_1, r_2) - \Delta_{12}^2 r_1^2 \Psi_1(r_1) - \Delta_{13}^2 r_2^2 \Psi_1(r_2)}{1 + 2\beta(\Delta_{12}s_1 + \Delta_{13}s_2) + \Delta_{12}\Delta_{13}\Phi_2(r_1, r_2) + \Delta_{12}^2 r_1^2 \Psi_2(r_1) + \Delta_{13}^2 r_2^2 \Psi_2(r_2)}, \\ \varepsilon_{\text{eff},yy} &= \varepsilon_1 \frac{1 - 2\beta(\Delta_{12}s_1 + \Delta_{13}s_2) + \Delta_{12}\Delta_{13}\Phi_2(r_1, r_2) + \Delta_{12}^2 r_1^2 \Psi_2(r_1) + \Delta_{13}^2 r_2^2 \Psi_2(r_2)}{1 + 2\alpha(\Delta_{12}s_1 + \Delta_{13}s_2) - \Delta_{12}\Delta_{13}\Phi_1(r_1, r_2) - \Delta_{12}^2 r_1^2 \Psi_1(r_1) - \Delta_{13}^2 r_2^2 \Psi_1(r_2)}. \end{aligned} \qquad (26)$$

Here, s_1 and s_2 are the concentrations of the inclusions of either sort (see formula (1)) and α and β are constants given by

$$\alpha = \frac{1}{\pi}(2 - A_1 - B_1), \quad \beta = \frac{1}{\pi}(2 + A_2 + B_2) \\ (\alpha + \beta = 1, \quad \alpha = 0.7047). \qquad (27)$$

The above formulas yield a number of particular results that are noteworthy.

(1) If the fibers have the same permittivity, $\varepsilon_2 = \varepsilon_3$ $(\Delta_{12} = \Delta_{13})$, expressions (26) are transformed to the form

$$\begin{aligned} \varepsilon_{\text{eff},xx} &= \varepsilon_1 \frac{1 - 2\alpha s\Delta_{12} - \Delta_{12}^2 \Omega_1(r_1, r_2)}{1 + 2\beta s\Delta_{12} + \Delta_{12}^2 \Omega_2(r_1, r_2)}, \\ \varepsilon_{\text{eff},yy} &= \varepsilon_1 \frac{1 - 2\beta s\Delta_{12} + \Delta_{12}^2 \Omega_2(r_1, r_2)}{1 + 2\alpha s\Delta_{12} - \Delta_{12}^2 \Omega_1(r_1, r_2)}. \end{aligned} \qquad (28)$$

Here, $s = s_1 + s_2$ is the total inclusion concentration and functions $\Omega_{1,\,2}(\cdot)$ are expressed as

$$\begin{aligned} \Omega_1(r_1, r_2) &= -\Phi_1(r_1, r_2) - r_1^2 \Psi_1(r_1) - r_2^2 \Psi_1(r_2), \\ \Omega_2(r_1, r_2) &= \Phi_2(r_1, r_2) + r_1^2 \Psi_2(r_1) + r_2^2 \Psi_2(r_2). \end{aligned} \qquad (29)$$

Based on formulas (28) and (29), one can conclude that the permittivity of the two-phase system remains anisotropic, since the radii of the cylindrical fibers arranged into rows alternating along the x axis are different.

(2) The two-phase material remains anisotropic even if fibers of one sort are present. Assuming $\varepsilon_3 = \varepsilon_1$ $(\Delta_{13} = 0)$ or $r_2 = 0$, which is the same thing, we obtain

$$\begin{aligned} \varepsilon_{\text{eff},xx} &= \varepsilon_1 \frac{1 - 2\alpha s_1\Delta_{12} - \Delta_{12}^2 r_1^2 \Psi_1(r_1)}{1 + 2\beta s_1\Delta_{12} + \Delta_{12}^2 r_1^2 \Psi_2(r_1)}, \\ \varepsilon_{\text{eff},yy} &= \varepsilon_1 \frac{1 - 2\beta s_1\Delta_{12} + \Delta_{12}^2 r_1^2 \Psi_2(r_1)}{1 + 2\alpha s_1\Delta_{12} - \Delta_{12}^2 r_1^2 \Psi_1(r_1)}. \end{aligned} \qquad (30)$$

In this case, the anisotropy of the inhomogeneous material is caused by the fact that the fibers are arranged into rows in the y-axis direction, which alternate with a period of $2h$ along the x axis. The spacings between neighboring inclusions in the longitudinal and transverse directions are different.

(3) If the parameters of the system are such that the equality

$$\Delta_{12}s_1 = -\Delta_{13}s_2 \qquad (31)$$

is satisfied, Eqs. (26) take the form

$$\begin{aligned} \varepsilon_{\text{eff},xx} &= \varepsilon_1 \frac{1 - \Delta_{12}^2 \Gamma_1(r_1, r_2)}{1 + \Delta_{12}^2 \Gamma_2(r_1, r_2)}, \\ \varepsilon_{\text{eff},yy} &= \varepsilon_1 \frac{1 + \Delta_{12}^2 \Gamma_2(r_1, r_2)}{1 - \Delta_{12}^2 \Gamma_1(r_1, r_2)}, \end{aligned} \qquad (32)$$

where

$$\begin{aligned} \Gamma_1(r_1, r_2) &= \frac{r_1^2}{r_2^2}\Phi_1(r_1, r_2) - r_1^2 \Phi_1(r_1) - r_1^2 \Psi_1(r_2), \\ \Gamma_2(r_1, r_2) &= -\frac{r_1^2}{r_2^2}\Phi_2(r_1, r_2) + r_1^2 \Phi_2(r_1) + r_1^2 \Psi_2(r_2). \end{aligned} \qquad (33)$$

As is seen, the components of the effective permittivity tensor do not contain first-order parameter Δ_{12} but still retain the anisotropic properties.

At equal concentrations of the phases $(s_1 = s_2)$, equality (31) implies that their permittivities are related as

$$\varepsilon_1 = \sqrt{\varepsilon_2 \varepsilon_3}. \qquad (34)$$

(4) If all the fibers have the same radii and permittivities, the two-phase material demonstrates isotropic properties because the inclusions are located in the vertices of the square cell (in this case, the same formulas for the effective permittivity are derived if a square of side $h/2$ is taken as a calculation cell). The electrical

performance of such a system has been studied in detail in numerous papers tracing back to Rayleigh's classical work [7].

CONCLUSIONS

The present-day theory of composite insulators covers largely two-phase isotropic materials in which the matrix contains one sort of inclusion. Such systems are easy to treat analytically and, therefore, have been studied most thoroughly. The insertion of two or more components allows for development of insulating materials with varied properties and structures. Investigation into multicomponent systems is still in its infancy.

The effective parameters of the material considered in this paper are calculated in the case when the interaction between the inclusions is described by only the first dipoles in an infinite set of dipoles with decreasing magnitudes of the moments. If an exact description of the system's properties is needed, higher order dipoles should be taken into account. Such a need arises, for example, when the inclusions are closely packed or when the difference between the permittivities of the matrix and inclusions is large. The following calculation scheme may then be suggested. The interaction between neighboring inclusions is considered in detail (i.e., a large number of image dipoles is taken into account), while the interplay of widely spaced inclusions is considered approximately (by taking into account several initial image dipoles or even one dipole). With such an approach, the mean parameters of the composite material still can be described analytically. The results of simulation corroborate the efficiency of the calculations performed. The fact is that, as the dipole order grows, the dipole moments decrease drastically, especially for widely spaced inclusions.

REFERENCES

1. Yu. N. Urzhumtsev, B. D. Annin, A. L. Kalashkarov, A. G. Kolpakov and V. Z. Parton, *Design of Composite Materials and Structure Elements* (Nauka, Novosibirsk, 1993) [in Russian].
2. N. S. Bakhvalov and G. P. Panasenko, *Averaging of Processes in Periodic Media* (Nauka, Moscow, 1984) [in Russian].
3. R. I. Nigmatulin, *Fundamentals of Mechanics of Heterogeneous Media* (Nauka, Moscow, 1978) [in Russian].
4. T. L. Chelidze, A. I. Derevyanko, and O. D. Kurilenko, *Electrical Spectroscopy of Heterogeneous Systems* (Naukova Dumka, Kiev, 1977) [in Russian].
5. Yu. P. Emets and Yu. P. Onofrishuk, IEEE Trans. Dielectr. Electr. Insul. **3**, 87 (1996).
6. Yu. P. Emets, Zh. Éksp. Teor. Fiz. **118**, 1207 (2000) [JETP **91**, 1046 (2000)].
7. Lord Rayleigh, Philos. Mag. **34**, 481 (1892).

Translated by M. Astrov

Technical Physics, Vol. 50, No. 2, 2005, pp. 213–216. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 73–77.
Original Russian Text Copyright © 2005 by Komolov, Aliaev, Potyupkin, Buzin.

SOLID-STATE ELECTRONICS

Electronic Structure at the PTCDA/GaAs and NTCDA/GaAs Interfaces

S. A. Komolov, Yu. G. Aliaev, N. V. Potyupkin, and I. S. Buzin

Fock Research Institute of Physics, St. Petersburg State University,
Ul'yanovskaya ul. 1, Petrodvorets, St. Petersburg, 198504 Russia

Received May 26, 2004

Abstract—The formation of the interface between the GaAs(100) single-crystal surface and PTCDA and NTCDA organic semiconductors is investigated. The method of total current spectroscopy makes it possible to trace the formation of the interfacial electronic structure. The two organic materials and the GaAs substrate are bonded together when the π electron cloud of an aromatic ring spreads toward the substrate. This modifies the electronic states of interfacial organic molecules and generates a dipole at the interface. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

Organic semiconductors have recently received widespread attention, and now they are finding increasing application, e.g., in various devices. In their structure and principle of operation, they may be direct analogues of the conventional (inorganic) semiconductor devices or have a radically new design (e.g., multilayer light-emitting devices [1]).

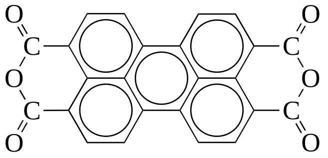
Electronic processes occurring both in the bulk of a semiconductor and at the contact of the semiconductor with a dissimilar material are to a large extent responsible for the operation of semiconductor devices. The contact properties are crucially dependent on phenomena taking place at the interface, a near-contact region with electronic properties distinct from those in the bulk due to interaction between the contacting materials. Electronic states at the interface may undergo considerable modification. In particular, new interfacial states may appear or the intrinsic states of contacting material change [2–4]. This may lead to electron charge transfer in a certain direction relative to the interface, which shows up as a bending of the energy band in the contacting materials and the polarization of organic molecules.

In this work, we study thin films of two well-known organic semiconductors, 1,4,5,8-naphthalene tetracarboxydianhydride (NTCDA) and 3,4,9,10-perylene tetracarboxydianhydride (PTCDA), that are formed on the GaAs(001) surface. While the interaction of organic molecules with metals and elementary semiconductors is well understood [5–8], the interface between organic semiconductors and binary semiconductors is poorly known [9]. The materials considered in this work (Fig. 1), polyaromatic compounds with extra carboxydianhydride groups, are the most extensively studied and widely used organic semiconductors.

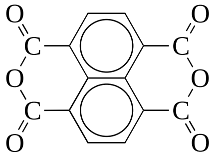
The method of total current spectroscopy (TCS) makes it possible to study the structure and density of unoccupied electronic states at the organic–inorganic interface (at various film thicknesses starting from submonolayer coverages) and the surface potential. In addition, this method is nondestructive, which is extremely important for organic molecules.

EXPERIMENTAL

The organic films were deposited *in situ* in a vacuum chamber maintained under an extralow pressure (5×10^{-8} Pa) during the deposition and spectral measurements. The chamber was equipped with a four-grid electron energy analyzer, which enabled us to apply solid surface analysis techniques such as TCS, Auger electron spectroscopy (AES), and low energy electron diffraction (LEED). In the TCS mode (the basic one used in the experiments), a parallel electron beam of given energy was directed onto the sample surface and the current in the sample circuit was recorded as a func-



1



2

Fig. 1. Chemical configuration of (*1*) PTCDA and (*2*) NTCDA molecules.

1063-7842/05/5002-0213$26.00 © 2005 Pleiades Publishing, Inc.

tion of the incident electron energy [10]. The energy of the probing beam was varied in the range 0–30 eV, and the current density was about 10^{-6} A/cm^2. The fine structure was resolved by recording the first derivative of the current with respect to energy ($S(E) = dI(E)/dE$) using incident electron energy modulation and a differentiating phase detector. The electron beam was directed normally to the surface and focused to a diameter of 0.2–0.4 mm.

As was found previously for the case of molecular solids, which NTCDA and PTCDA films refer to, the fine structure of the total current spectrum reflects the density of unoccupied electronic states above the vacuum level [11]. If the energy of incident electrons corresponds to the allowed band of the sample material, the coefficient of electron transfer from the vacuum to the sample increases and, as a result, so does the current in the sample circuit. When the energy of incident electrons becomes equal to the energy gap, the coefficient of electron transfer and the current decline. For organic molecules, the electronic states of solids studied by TCS are the π^* and σ^* orbitals of the molecules. With the TCS method, one can also keep track of the surface potential of the sample [10].

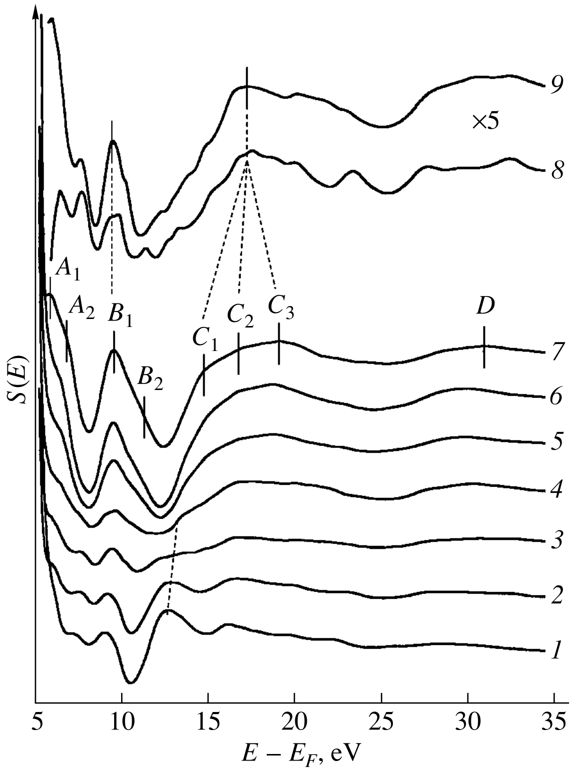


Fig. 2. Evolution of the total current spectra in the course of PTCDA film deposition: (*1*) total current spectrum of the pure GaAs substrate, (*2*–*6*) variation of the spectrum with the film thickness, (*7*) total current spectrum of the PTCDA film formed, and (*8*, *9*) difference curves (curves *8* and *9* are obtained by subtraction of the substrate spectrum from spectra *2* and *3*, respectively). The difference curves are enlarged fivefold.

Before deposition, the sample surface was specially conditioned. Prior to being placed into the chamber, the GaAs single crystal was first chemically etched in HF to remove the oxide film and then rinsed in distilled water. However, the Auger spectra taken from the surface thus prepared showed that it is heavily contaminated by carboniferous impurities. Therefore, the surface was cleaned further using ion etching and subsequent annealing until the carbon peaks in the Auger spectra completely disappeared. In the following experiments, the secondary cleaning of the substrate was performed by ion etching. The total current spectrum of the surface cleaned in such a manner was in good agreement with the well-known spectra for the single-crystal GaAs(001) surface [12]. It should be noted that the composition of the surface subjected to different actions (deposition and thermal desorption of organic molecules, ion etching, and heating) varied and somewhat differed from the stoichiometric one. This variation caused corresponding variations in the total current spectrum of the clean GaAs substrate (though only the relative intensities of different spectral features varied, with their energy positions remaining unchanged) and in the work function.

The films were grown by thermal sublimation of the organic molecules followed by deposition onto the substrate. The configuration of NTCDA and PTCDA molecules is shown in Fig. 1. (For details of growing PTCDA and NTCDA films, see [11, 13, 14].) Our experimental setup made it possible to record the total current spectra *in situ*, i.e., immediately during the film growth.

In addition to the molecular film growth experiments, experiments aimed at removing the films by thermal desorption were carried out. To this end, the structure was heating to a certain temperature and then the total current spectrum was measured. Series of such measurements with the successively increasing heating temperature were performed for either type of molecule.

RESULTS AND DISCUSSION

The families of the spectra recorded as the film thickness was successively increased are presented in Fig. 2 for the PTCDA (curves *1*–*7*) and in Fig. 3 for the NTCDA (curves *1*–*9*). Both spectra vary with the film thickness in the same manner: namely, the spectral features characteristic of the organic compounds continuously grow, while those characteristic of the substrate decay. When the film thickness reaches a certain value, the total current spectrum stops varying and takes the form typical of the volume phase of the compounds. The PTCDA spectrum has the following main features: A_1 (5.8 eV), A_2 (6.9 eV), B_1 (9.5 eV), B_2 (11.3 eV), C_1 (14.7 eV), C_2 (16.7 eV), C_3 (19.1 eV), and D (31 eV)

(Fig. 2, curve *7*). This spectrum was in good agreement with the spectra taken from PTCDA films deposited on other substrates [13, 14]. The basic features in the NTCDA spectrum were *A* (6.3 eV), B_1 (8.5 eV), B_2 (9.9 eV), *C* (12.4 eV), *D* (16.1 eV), *E* (22.6 eV), and *F* (28.8 eV) (Fig. 3, curve *9*). This spectrum also agrees well with the previous results [14].

To gain a better insight into the interaction of the organic molecules with the substrate, the spectra were processed as follows. From the spectrum taken at a certain thickness of the film, the appropriately scaled spectrum of the substrate was subtracted. The scaling procedure takes into account that the contribution from the substrate drops with increasing film thickness [3, 4]. The scaling factor was calculated from that feature in the total current spectrum of the substrate which does not overlap with a spectral feature of the deposit. Then, the intensity of this feature was measured both in the spectrum of the film having a certain thickness and in the spectrum of the pure substrate. The ratio of these two values was used as the scaling factor at a given thickness of the film. In order to increase the accuracy of finding the scaling factor, it was calculated for several features and the results obtained were averaged.

The difference curves are presented in Fig. 2 for the PTCDA (curves *8*, *9*) and in Fig. 3 for the NTCDA (curves *10–12*). It is seen that they, in general, are similar to the total current spectra of the compounds. In particular, the difference curves for both compounds contain the same features as the NTCDA and PTCDA spectra. Moreover, the similarity between the difference curves and the spectra becomes more evident at higher energies (E > 14 eV for the NTCDA and E > 20 eV for the PTCDA, respectively). However, there are also noticeable discrepancies between the difference curves and the spectra of the compounds. For example, feature B_2 is missing in the case of PTCDA (Fig. 2) and maxima C_1–C_3 merge into one maximum. The same is basically true for the NTCDA curves: one maximum instead of B_1 and B_2 and weaker maximum *C*.

Such dissimilarities between the interfacial total current spectra and the spectra taken from the volume phase of the materials under study (the latter characterize the materials as such) indicate that the film–substrate interaction alters the molecular orbital configuration. In the case of PTCDA/GaAs and NTCDA/GaAs, several features in the interfacial total current spectra merge into one and the spectra shift along the energy axis (relative to the total spectra). This testifies that the interface electronic states are the modified electronic states of the initial macromolecules. At the same time, interface states that are untypical of the macromolecules do not form, as distinct from the case when such films are deposited on the copper substrate [5].

At low energies (less than 8 eV), two circumstances regarding the difference curves should be taken into account. First, since the beginning of the spectrum involves the decay of the primary peak [10], whose position may vary with the surface potential, subtraction of one spectrum from another may introduce a great error. Second, the free path of electrons in solids to a great extent depends on their energy [15], increasing drastically at low energies. Therefore, the signal from the low-energy features of the deposit is visualized at larger thicknesses as compared with the high-energy features. Accordingly, the error in the low-energy part of the difference curves increases considerably and the curves themselves become irregular.

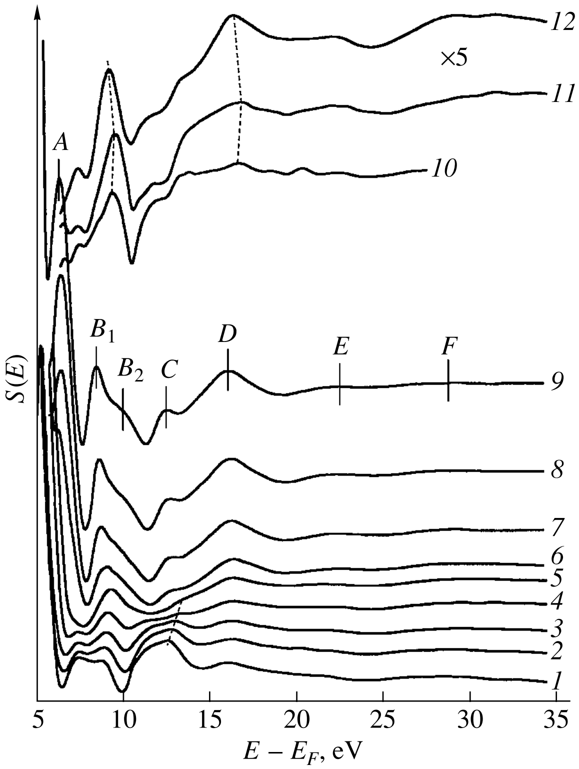


Fig. 3. Evolution of the total current spectra in the course of NTCDA film deposition: (*1*) total current spectrum of the pure GaAs substrate, (*2–8*) variation of the spectrum with the film thickness, (*9*) total current spectrum of the NTCDA film formed, and (*10–12*) difference curves (curves *10, 11,* and *12* are obtained by subtraction of the substrate spectrum from spectra *2, 3,* and *4,* respectively). The difference curves are enlarged fivefold.

There is another trend in the spectra of the PTCDA/GaAs and NTCDA/GaAs interfaces early in their formation. The spectral features characteristic of the substrate shift toward higher energies as the concentration of the molecules on the surface increases, the total shift between the curve from the pure substrate and the final curve where the GaAs features still persist being about 0.4 eV. Such a shift is due to the bending of bands in the substrate, and the direction of the shift suggests that a negative charge is transferred to the semiconductor surface. Early in the interface formation, the work function first decreases slightly (by 0.1–0.2 eV) and then starts to increase. By the time the band bending in the substrate reaches 0.4 eV, the work function approaches the value typical of the substrate. This is because molecules on the surface take a positive charge and a dipole arising between the molecules and the substrate surface compensates for the change in the work function. As for the spectral features typical of the molecules deposited, their energy positions somewhat fluctuate early in the interface formation (the fluctuations are most distinctly seen in the difference curves for the NTCDA, Fig. 3). The fluctuations are likely to reflect the fact that both competing processes described above (namely, the negative charging of the substrate surface and the positive charging of the molecules deposited) occur simultaneously. As the film grows further, the interfacial dipole is screened and the work function of the sample becomes equal to the value typical of the deposit.

Thus, physically, the formation of the electronic structure at the PTCDA/GaAs and NTCDA/GaAs interfaces exhibits two specific features: (i) the electron density is transferred from the molecules to the substrate and (ii) the deposit–substrate interaction involves electron states with energies of up to 11 eV. This energy range is occupied largely by the π^*-orbitals of aromatic rings of the molecules [11]. It follows from the aforesaid that the basic mechanism of interaction of NTCDA and PTCDA molecules with the GaAs surface is spreading of the π electron cloud of the aromatic rings from the molecules to the substrate. The thickness of the interface was roughly estimated as 3 nm, which corresponds to two to three monolayers for both PTCDA and NTCDA. This value is smaller than the thickness of the copper–deposit interface [2], where the molecules interact with the surface via oxygen-to-copper bonding.

The experiments where the macromolecules deposited were thermally desorbed from the GaAs surface indicate that the molecules partially decompose rather than being completely desorbed. As a result, features typical of both pure GaAs and amorphous carbon [16] appear in the total current spectrum even after high-temperature (750 K) annealing. NTCDA molecules are desorbed from the surface more readily than PTCDA molecules. PTCDA–substrate bonds are stronger than NTCDA–substrate bonds, possibly because of a larger size of the aromatic ring and, hence, a larger number of bond-forming π electrons.

CONCLUSIONS

The formation of the interface between organic macromolecular NTCDA and PTCDA films and the GaAs(100) substrate was investigated. Using the method of total current spectroscopy, we kept track of the occurrence of unoccupied electronic states by increasing the film thickness starting from submonolayer coverages. It was found that molecule-to-substrate bonding, which is due to π electrons of the aromatic rings of the molecules, modifies π and π^* electronic states of the molecules at the interface. Because of the larger number of π electrons in a PTCDA molecule, it is bonded to the substrate more strongly. Therefore, PTCDA molecules partially decompose during annealing in the desorption experiments.

ACKNOWLEDGMENTS

The authors thank A. Sidorenko (Institut für Physicaliche Chemie, Universität Tuebingen) for submitting the single-crystal GaAs sample.

This study was supported by the Russian Foundation for Basic Research (project no. 02-03-32751) and the Russian State Program “Surface Atomic Structures.”

REFERENCES

1. S. R. Forrest, Chem. Rev. **97**, 1793 (1997).
2. S. A. Komolov and Y. G. Aliaev, Colloids Surf., A **239**, 55 (2004).
3. S. A. Komolov and Y. G. Aliaev, Phys. Low-Dimens. Semicond. Struct. **9–10**, 55 (2002).
4. S. A. Komolov and Y. G. Aliaev, Phys. Low-Dimens. Semicond. Struct. **5–6**, 87 (2003).
5. S. A. Komolov, N. B. Gerasimova, Y. G. Aliaev, *et al.*, Phys. Low-Dimens. Semicond. Struct. **1–2**, 253 (2002).
6. Y. Hirose, C. I. Wu, V. Aristov, *et al.*, Appl. Surf. Sci. **113–114**, 291 (1997).
7. D. Gador, C. Buchberger, R. Fink, *et al.*, J. Electron Spectrosc. Relat. Phenom. **96**, 11 (1998).
8. I. G. Hill, J. Schwartz, and A. Kahn, Organ. Electron. **1**, 5 (2000).
9. Y. Hirose, W. Chen, E. I. Haskal, *et al.*, Appl. Phys. Lett. **64**, 3482 (1994).
10. S. A. Komolov, *Integral Secondary-Electron Spectroscopy of Surface* (LGU, Leningrad, 1986) [in Russian].
11. N. B. Gerasimova, S. A. Komolov, Y. G. Aliaev, *et al.*, Phys. Low-Dimens. Semicond. Struct. **1–2**, 119 (2001).
12. S. A. Komolov and Ya. Galat, Vestn. Leningr. Univ., No. 4, 24 (1982).
13. S. A. Komolov, N. B. Gerasimova, A. G. Sidorenko, *et al.*, Pis’ma Zh. Tekh. Fiz. **26** (24), 80 (2000) [Tech. Phys. Lett. **26**, 1110 (2000)].
14. S. A. Komolov, N. B. Gerasimova, A. G. Sidorenko, *et al.*, Phys. Low-Dimens. Semicond. Struct. **7–8**, 93 (2001).
15. S. A. Komolov, E. F. Lazneva, A. S. Komolov, *et al.*, Pis’ma Zh. Tekh. Fiz. **29** (23), 13 (2003) [Tech. Phys. Lett. **29**, 974 (2003)].
16. S. A. Komolov, N. B. Gerasimova, and A. O. Morozov, Phys. Low-Dimens. Semicond. Struct. **10**, 35 (1997).

Translated by M. Lebedev

Technical Physics, Vol. 50, No. 2, 2005, pp. 217–219. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 78–81.

Process of Fabricating Semiconductor Microcavities and Photon Crystals

E. M. Arakcheeva, A. V. Nashchekin, V. A. Solov'ev, E. M. Tanklevskaya, M. V. Maksimov, S. G. Konnikov, S. A. Gurevich, and N. N. Ledentsov

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: kathy.quantum@mail.ioffe.ru

Received June 23, 2004

Abstract—A process of fabricating microcavities and photon crystals in GaAs structures by means of electron lithography and reactive ion etching is described. Two types of structures, with micropillars and with photon crystals, are considered. The latter structures have the form of a square or hexagonal array of holes in a planar waveguiding structure. The minimal diameter of the micropillars is 100 nm, and their height is 700 nm. The size of the holes in the photon crystals and the photon crystal period are controllably varied from 140 to 500 nm and from 400 to 1000 nm, respectively. The etch depth of the crystals is more than 350 nm. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Semiconductor cavities are of both scientific and applied interest, since they are viewed as promising elements for optoelectronic devices. The semiconductor cavity is a structure in which the electromagnetic field is quantized in one, two, or three directions. The electromagnetic field in the microcavity can be confined using distributed Bragg reflectors or by means of total internal reflection. Specifically, in a three-dimensional microcavity, which is referred to as a micropillar, the light is confined in the vertical direction by using upper and lower distributed Bragg mirrors; in the horizontal direction, due to total internal reflection from the air–semiconductor interface. Using microcavities, one can control the intensity of light–material interaction via increasing or decreasing the overlap between electromagnetic field modes and/or between the electron wave functions. It has been shown theoretically [1] that a radiator placed in a microcavity may exhibit enhanced spontaneous recombination (the Purcell effect). At present, microcavities are being widely used in advanced optoelectronic devices, such as vertical microcavity surface-emitting lasers and resonance light-emitting diodes.

Quantum dots (QDs), which are produced by the method of self-organization in the process of growth, offer a number of properties that make them very promising as an active medium of semiconductor cavities. Since carriers are spatially confined in an ideal quantum dot, its electron spectrum consists of a number of discrete levels separated by forbidden gaps and resembles the electron spectrum of an atom. The small width of the radiation lines of a single QD makes it possible to observe the Purcell effect in a semiconductor microcavity [2]. It seems very attractive to place quantum dots into microcavities to create sources of single photons [3].

Development of next-generation lasers and light-emitting diodes based on photon crystals is another intriguing field of research [4]. The photon crystal is a semiconductor or insulating periodic structure with a period comparable to the radiation wavelength. In this medium, photons form a band structure that is similar to the band structure of electrons in a solid. These crystals may serve as very efficient one-dimensional or two-dimensional distributed Bragg mirrors. In a microcavity surrounded by such a distributed two-dimensional mirror, the radiation of the active material may be totally concentrated into one allowed mode. It is expected that such a possibility will help to design a light-emitting device that combines the advantages of a laser (coherent and weakly divergent radiation) and light-emitting diode (the absence of a threshold) [5]. Photon crystals with quantum wells or quantum dots may be used as active elements, e.g., lasers or light-emitting diodes [6]. Optical integrated circuits based on microcavities and photon crystals are currently under development with the aim of creating a quantum computer.

MICROCAVITY FABRICATION TECHNOLOGY

To create sources of radiation that are based on the Purcell effect and also sources of single photons, small-size microcavities (it is best if they contain only one QD) are necessary. A typical average spacing between InAs/GaAs QDs produced by the self-organization technique in the process of growth is about 100 nm.

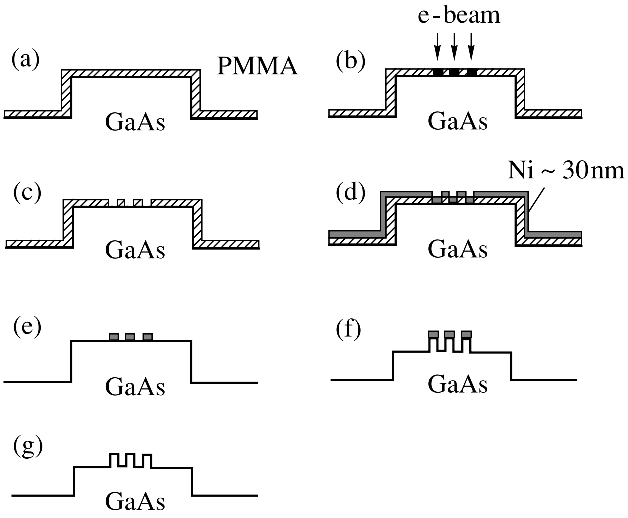


Fig. 1. Microcavity fabrication process: (a) application of a 350-nm-thick PMMA layer, (b) electron-beam lithography, (c) PMMA development, (d) evaporation of a 30-nm-thick nickel layer, (e) PMMA liftoff, (f) reactive ion etching of the semiconductor, and (g) nickel mask stripping.

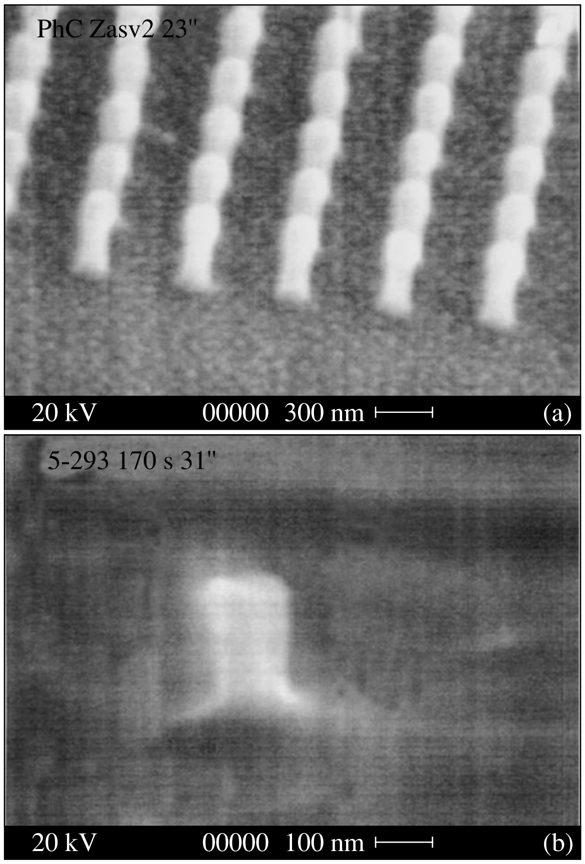


Fig. 2. SEM image of (a) the nickel-mask-covered array of mesas 140 nm in diameter and (b) a single mesa 140 nm in diameter.

This value determines the desired diameter of the micropillar.

The micropillars were made by optical and electron-beam lithography, as well as by chemical and reactive etching. First, a mask of AZ 5214 organic resist was applied on the surface of the structure by optical lithography. Through this mask, the whole structure except for 100 × 100-μm squares several millimeters apart was etched off. Then, mesa arrays were formed in these squares. The absence of the active material in the regions between the mesa arrays excludes photoluminescence from these regions under optical measurements.

Further, the sample was covered by a positive electron resist (≈350-nm-thick polymethylmethacrylate (PMMA) film, Fig. 1a) and electron-beam lithography was performed on a CamScan Series 4-88 DV 100 scanning electron microscope (Fig. 1b). The beam was controlled by a high-resolution eight-channel digital card built around a 16-bit DAC (ADLINK PCI-6208V). We devised a Delphy 5-written software that applies a voltage to the deflection system of the electron gun in 150-μV steps and also controls the exposure time per feature by applying voltage pulses of desired duration to the electron beam modulator. The lithography conditions were optimized by patterning groups of features with different periods (from 0.35 to 10 μm) and different exposure times (from 50 to 1000 μs depending on the proximity of adjacent features). The optimum beam current was 10–30 pA at an accelerating voltage of 15 kV.

Since PMMA is not plasma-resistant, it cannot be used as a mask for reactive ion etching. Therefore, after developing the PMMA film (Fig. 1c), we evaporated a ≈30-nm-thick nickel layer (Fig. 1d) and lifted off the PMMA in dimethylformamide (Fig. 1e). The nickel layer that remained on the surface after the liftoff served as a mask in subsequent reactive ion etching.

In this work, reactive ion etching was carried out in an RDE-300 (Alcatel, France) computerized diode-type rf plasma etcher. The etching parameters were chosen such that the lateral surfaces of the structure are vertical and contain a small number of etching-induced defects. The etching conditions were the following (Fig. 1f): the gas components were taken in the ratio $Cl_2 : BCl_3 : Ar = 1 : 4 : 16$ sccm, the gas pressure in the reactor was 1 Pa (the reactor was preevacuated to a pressure of 5×10^{-4} Pa), and the self-bias voltage was 200 V. Under these conditions, the average etching rate was ≈90 nm/min and the etch depth was 0.5–1.5 μm, depending on the parameters of the structure. The micropillar diameter-to-height ratio was 1 : 6, with the anisotropy of the wall remaining high. At the final stage of the process, the nickel mask left on the micropillar surfaces was removed by chemical etching (Fig. 1g). Figure 2 shows the SEM (CamScan) images of an array of mesas 140 μm in diameter and a single mesa 140 μm in diameter with the nickel mask intact.

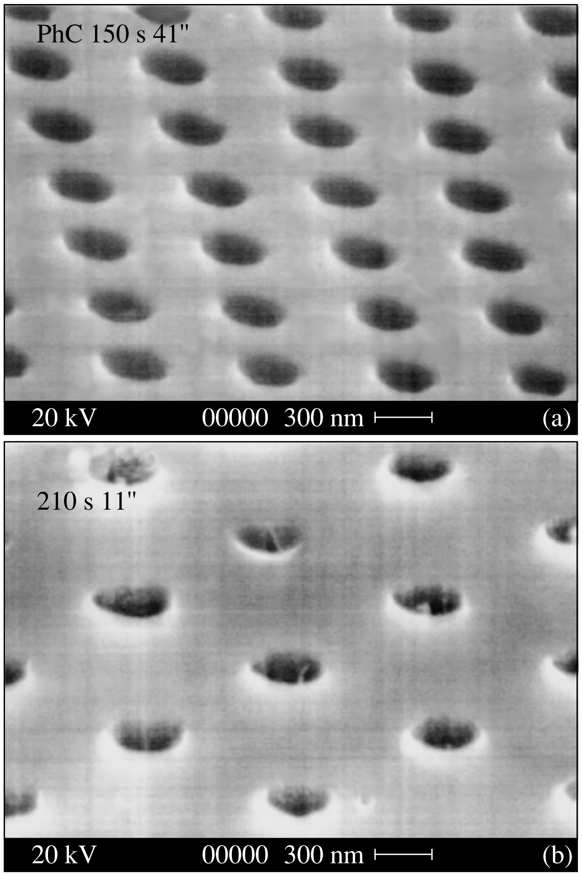


Fig. 3. SEM images of the (a) square and (b) hexagonal photon crystals.

PROCESS OF FABRICATING TWO-DIMENSIONAL PHOTON CRYSTALS

Since production of three-dimensional semiconductor photon crystals is right now a bugaboo, two-dimensional photon crystals have gained a wider application. Such crystals represent a two-dimensional square or hexagonal array of holes in a semiconductor structure. Two-dimensional photon crystals made in an epitaxial waveguiding structure are the most promising for application in optoelectronics. Here, the electromagnetic wave is confined in the direction parallel to the hole axes due to the waveguide effect. It has been shown that a hexagonal photon crystal has a photon band forbidden for the TE mode [7].

The technology of two-dimensional semiconductor photon crystals is in many respects similar to the microcavity fabrication process illustrated in Fig. 1. At the same time, the two technologies somewhat differ. In the former case, a 20-nm-thick nickel layer is evaporated on the surface before PMMA application. Then, a positive PMMA resist is applied and electron-beam lithography is accomplished. After the PMMA layer has been developed, the sample is etched with an argon beam. The difference in the rates of Ar etching for the nickel and PMMA layers is not too large. Therefore, the nickel layer could be completely removed on those areas where the PMMA is absent. A mask for subsequent reactive ion etching is thus formed. The etch parameters for the photon crystals were chosen so as to make the walls vertical and smooth and also to minimize the concentration of nonradiative recombination centers on the semiconductor surface. Figures 3a and 3b show the images of the prepared photon crystals with square and hexagonal arrays of holes. The period of the “square” photon crystal is 660 nm, and the hole diameter is 320 nm. In the “hexagonal” photon crystal, the hole spacing and hole diameter are 800 and 420 nm, respectively. The holes are circular and have even edges. The examination of the cleaved surface of the crystals showed that the etch depth equals 350 nm and the walls are vertical and smooth.

CONCLUSIONS

A process of fabricating AlGaAs/GaAs-based semiconductor microcavities and photon crystals by electron lithography and reactive ion etching is described. The minimum mesa size is about 100 nm, which allowed us to fabricate single-quantum-dot microcavities. The hole spacing and the hole diameter in the photon crystals vary from 400 to 1000 nm and from 140 to 500 nm, respectively. This corresponds to the wavelength range 800–1300 nm, which is now common in fiber-optic communication lines. The tentative optical study of mesas and photon crystals the active region of which contains quantum dots suggests that reactive ion etching degrades the optical grade of the material insignificantly.

ACKNOWLEDGMENTS

This work was supported by the program “Advanced Materials and Structures” (Physical Science Department, Russian Academy of Sciences) and the program “Low-Dimensional Quantum Structures” (the Presidium of the Russian Academy of Sciences).

REFERENCES

1. E. M. Purcell, Phys. Rev. **69**, 681 (1946).
2. J. M. Gérard, B. Sermage, B. Gayral, *et al.*, Phys. Rev. Lett. **81**, 1110 (1998).
3. Zhiliang Yuan, B. E. Kardynal, R. M. Stevenson, *et al.*, Science **295**, 102 (2002).
4. *Microcavities and Photonic Bandgaps: Physics and Application*, Ed. by C. Weisbuch and J. Rarity (Kluwer, Dordrechet, 1996), NATO ASI Ser., Ser. E Vol. 324.
5. R. De La Rue and Ch. Smith, Nature **408**, 653 (2000).
6. A. A. Erchak, D. J. Ripin, Sh. Fan, *et al.*, Appl. Phys. Lett. **78**, 563 (2001).
7. H. Benisty, C. Weisbuch, D. Labilloy, *et al.*, J. Lightwave Technol. **17**, 2063 (1999).

Translated by A. Khzmalyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 220–226. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 82–88.

OPTICS, QUANTUM ELECTRONICS

Interference of Coherently Scattered Signals in the Laser SRS–CARS Monitoring of Hydrogen Concentration in Gas Mixtures

G. M. Mikheev, D. G. Kaluzhnyĭ, A. Yu. Popov, and T. N. Mogileva
Institute of Applied Mechanics, Ural Division, Russian Academy of Sciences, Izhevsk, 426067 Russia
e-mail: gmmikheev@udmnet.ru

Received March 23, 2004

Abstract—The possibility of measuring the hydrogen impurity concentration in dense gas mixtures by coherent anti-Stokes Raman scattering (CARS) is studied. In this technique, biharmonic laser pumping based on stimulated Raman scattering (SRS) in compressed hydrogen is used. Because of the interference between the coherent scattering components from buffer gas molecules and molecules of the impurity to be detected, the signal recorded may depend on the hydrogen concentration by a parabolic law, which has a minimum and makes the results uncertain. It is shown that this uncertainty can be removed if the frequency of the biharmonic laser pump, which is produced by the SRS oscillator, somewhat differs from the frequency of molecular oscillations of hydrogen in the test mixture. A sensitivity of 5 ppm is obtained as applied to the hydrogen–air mixture under normal pressure. The coefficients of the nonresonance cubic hyperpolarizability of molecules and atoms are measured relative to nitrogen in a number of gas media. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

Selective and rapid diagnostics of hydrogen in condensed media is presently a challenge in applications [1–6]. Also, in physical experiments on the interaction of hydrogen with metals, alloys, and semiconductors, there often appears the need for express monitoring of molecular hydrogen in gas mixtures [7]. Such experiments are aimed, for example, at finding ways for preventing tritium leakage through the thermonuclear reactor walls or at devising coatings to protect steel from hydrogenation, embrittlement, and fracture. Usually, such experiments determine the nonstationary flow rate of hydrogen, which diffuses through a membrane to a vacuum chamber, with a gage at a constant evacuation rate of a vacuum pump. In another simple method, hydrogen bubbles on the exit side of a membrane that is in contact with glycerin or alcohol are observed through a microscope. Clearly, these, as well as the other available methods (such as mass-spectrometric and chromatographic methods and also the unique selective method, which evaluates the hydrogen content in a gas mixture by measuring the deflection of a probing laser beam from the surface of a passive palladium film sensor heated by the modulated radiation of an argon laser [2]), are finding limited application and do not allow researchers to perform many physical experiments where real-time remote selective monitoring of hydrogen in gas mixtures is required. Therefore, for these and some other experiments (i.e., those that study hydrogen emission from metals during pulsed laser melting [8]), a simplified version of the CARS method where biharmonic laser pumping (BLP) is induced by SRS in compressed hydrogen [9–12] is promising. In this regard, it is of interest to further refine this nonlinear optical SRS–CARS method of hydrogen diagnostics in gas mixtures. It should be noted that the SRS–CARS method may become attractive for diagnosing other gases owing to hollow microstructured optical waveguides that have been recently developed [13], which open unique possibilities in nonlinear optics and substance spectroscopy [14, 15].

The purpose of this work is to study the effect of interference between nonlinear susceptibilities on SRS–CARS monitoring of the hydrogen concentration in dense gas mixtures.

1. THEORY

CARS is a four-photon parametric process in which two laser beams of frequencies ω_p and ω_s are mixed in a medium characterized by cubic nonlinear susceptibility $\chi^{(3)}$. As a result of the mixing, a coherent directional radiation at the anti-Stokes frequency $\omega_a = 2\omega_p - \omega_s$ is generated [16, 17]. In SRS–CARS hydrogen diagnostics, the medium is probed by biharmonic laser pumping at frequencies ω_p and ω_s, which meet the condition of approximate resonance

$$\omega_p - \omega_s \approx \Omega_{H2-BG}, \quad (1)$$

where Ω_{H2-BG} is the frequency of the Q_{01} (1) vibrational transition of the hydrogen that is in a hydrogen–buffer gas mixture of density ρ_{H2-BG} (Fig. 1).

The biharmonic laser pumping is produced in an SRS oscillator by focusing a high-power monochromatic laser radiation at a frequency ω_p into a cell containing compressed hydrogen at a pressure P_{H2} ($\omega_p - \omega_s = \Omega_{H2-H2}$, $\Omega_{H2-H2} < \Omega_0$, where Ω_{H2-H2} and Ω_0 are the frequencies of the Q_{01} (1) vibrational transition in hydrogen at pressure P_{H2} and at the gas pressure reduced to zero, respectively). The intensity I_a of scattered radiation at frequency ω_a is given by the relationship

$$I_a \sim \left|\chi^{(3)R} + \chi^{(3)NR}\right|^2 I_p^2 I_s, \tag{2}$$

where I_p and I_s are the radiation intensities at frequencies ω_p and ω_s, respectively; $\chi^{(3)R} = n_{H2}\gamma_r$ is the cubic resonance susceptibility of gas molecules under study; $\chi^{(3)NR} = n_{BG}\gamma_{BG}$ is the cubic nonresonance susceptibility that is due to electrons mostly of buffer gas molecules participating in a scattering event; γ_r and γ_{BG} are the cubic hyperpolarizabilities of the gas being detected and buffer gas (BG), respectively; and n_{H2} and n_{BG} are the concentrations of impurity hydrogen in the mixture and of buffer gas molecules, respectively (it is supposed that $n_{H2} \ll n_{BG}$). Cubic resonance susceptibility $\chi^{(3)R}$ is given by [16]

$$\chi^{(3)R} = \frac{1}{3}\Delta_k^n \frac{2\pi n_{H2} c^4}{h\Gamma\omega_s^4}\frac{d\sigma}{do}\frac{\Gamma}{\Omega_{H2-BG} - (\omega_p - \omega_s) - i\Gamma}, \tag{3}$$

where Γ is the half-width at half maximum of the Raman transition line, Δ_k^n is the difference between the populations of the levels, and $d\sigma/do$ is the molecular cross section of spontaneous Raman scattering in this transition.

As follows from expressions (2) and (3), when $\rho_{H2-BG} \ll 1$ Amagat unit (a rarefied gas mixture), nonresonance contribution $\chi^{(3)NR}$ may be neglected and the value of $[I_a/I_p^2 I_s]^{1/2}$ in expression (2) linearly depends on hydrogen concentration n_{H2}. However, when a low hydrogen concentration in a buffer gas is measured at atmospheric pressure (nonresonance contribution $\chi^{(3)NR}$ is significant), the dependence of $[I_a/I_p^2 I_s]^{1/2}$ on n_{H2} may become nonlinear because of interference between the nonlinear susceptibilities. Since the CARS process is coherent, the contributions from molecules of different types to the scattered signal intensity interfere instead of being added up [16]. This circumstance is noteworthy when monitoring hydrogen impurity concentrations in gases. As far as we know, this issue as applied to SRS–CARS diagnostics of hydrogen in gas mixtures has not been touched upon by other authors.

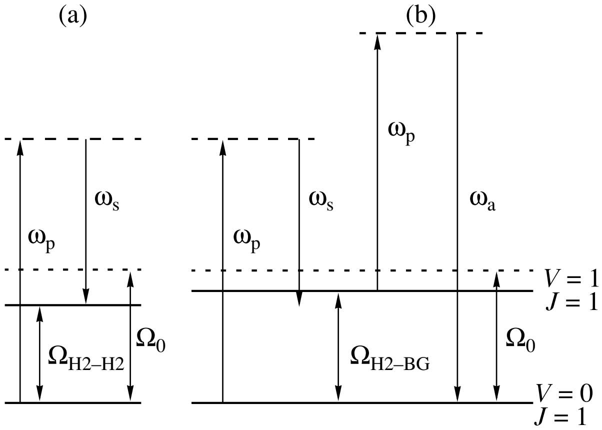


Fig. 1. Energy levels of the Q_{01} (1) Raman vibrational transition in hydrogen that is in (a) the SRS oscillator and (b) a rarefied gas mixture, as well as the quantum diagrams that illustrate the (a) SRS and (b) CARS methods.

In view of expression (3), formula (2) can be recast as [18]

$$\frac{I_a}{I_p^2 I_s} = b\left(\frac{\gamma^2}{1+\Delta^2}n_{H2}^2 - \frac{2n_{BG}\gamma\gamma_{BG}\Delta}{1+\Delta^2}n_{H2} + n_{BG}^2\gamma_{BG}^2\right), \tag{4}$$

where b is a dimensional coefficient,

$$\gamma = \frac{1}{3}\Delta_k^n \frac{2\pi c^4}{h\Gamma\omega_s^4}\frac{d\sigma}{do}, \tag{5}$$

$$\Delta = \frac{\Omega_{H2-H2} - \Omega_{H2-BG}}{\Gamma}. \tag{6}$$

As follows from formula (4), $I_a/I_p^2 I_s$ is a monotonically increasing function of n_{H2} at $\Delta \le 0$. However, if $\Delta > 0$, function (4) has a minimum if the impurity concentration is

$$n_{H2}^{\min} = \frac{n_{BG}\gamma_{BG}}{\gamma}\Delta. \tag{7}$$

In the absence of impurity molecules, the SRS–CARS signal is determined by nonresonance scattering by buffer gas particles,

$$\frac{I_a}{I_p^2 I_s} = b(n_{BG}^2\gamma_{BG}^2). \tag{8}$$

At $\Delta > 0$, introduction of hydrogen to a buffer gas decreases the signal, whose minimum is found from the expression

$$\left(\frac{I_a}{I_p^2 I_s}\right)_{\min} = b\left(\frac{n_{BG}^2\gamma_{BG}^2}{1+\Delta^2}\right). \tag{9}$$

As the hydrogen concentration increases further, the signal monotonically grows. Thus, at $\Delta > 0$, analytical results are uncertain, since, at low impurity concentra-

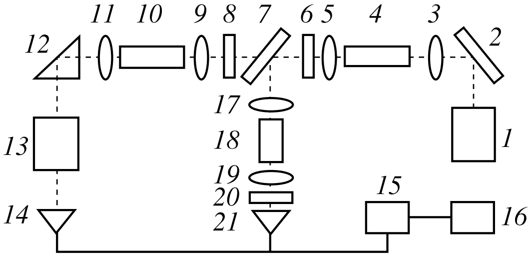


Fig. 2. Optical scheme of the experiment.

tions ($n_{H2} < 2n_{H2}^{min}$), the same signal is produced by two hydrogen concentrations. This uncertainty is eliminated when $\Delta \le 0$.

It should be noted that, according to (8), the slope of the experimental $[I_a/I_p^2 I_s]^{1/2}$ versus n_{BG} dependence is proportional to the nonresonance hyperpolarizability γ_{BG} of buffer gas particles. This fact allows ratios between the cubic hyperpolarizability coefficients of different gas particles to be determined.

2. EXPERIMENTAL TECHNIQUE

The experiments were carried out following the optical scheme for SRS–CARS hydrogen diagnostics described in [12]. A simplified version of the experimental scheme is illustrated in Fig. 2. As pump oscillator *1*, we used a single-wavelength passively Q-switched (with the help of a LiF : F_2^- crystal) YAG : Nd^{3+} laser. The polarization extraction of the radiation was accomplished with a one-active-element resonator [19]. The radiation of this laser was converted to the second harmonic with the help of a KTP crystal. The peak energy at the wavelength λ_p = 532 nm was 40 mJ; the FWHM τ_p, 16 ns; and the beam divergence, about 0.6 mrad. Deflecting mirror *2* and lens *3* (F_1 = 0.66 m) focused the radiation onto SRS cell *4* (L_1 = 0.86 m) filled with compressed molecular hydrogen at a pressure of 4 bar (T = 295 K). BLP generated in cell *4* as a result of SRS at the Q_{01} (1) vibrational transition was collimated by lens *5*, separated out from other SRS components with filter *6*, and focused onto measuring cell *10* (L_2 = 0.21 m) by objective lens *9* (F_2 = 0.1 m). The BLP intensity was attenuated, if necessary, by set of filters *8*. The anti-Stokes component arising in cell *10* as a result of one-dimensional CARS was directed to the entrance of monochromator *13* by lens *11* and prism *12*. Then, the signal was detected by photomultiplier *14* and multichannel system *15* of laser energy detection, which was linked to IBM PC *16* [7]. After being split by beam splitter *7*, part of the BLP was focused by lens *17* onto the center of reference cell *18* kept at a constant pressure of 4 bar. The anti-Stokes scattering component arising in reference cell *18* was frequency-selected by set of filters *20* and directed to photodiode *21* by lens *19*. This additional optical path (elements *17–21*) was used to normalize the signal generated in measuring cell *10* and to eliminate the effect of SRS-exciting laser intensity fluctuations. It should be noted that, essentially, reference optical path *17–21* allows one to measure the SRS–CARS signal in the measuring cell, which is proportional to $(I_a/I_p^2 I_s)^{1/2}$, without measuring the BLP component intensities.

Prior to performing the experiments, the measuring cell was evacuated to a pressure of less than 0.1 kPa. Then, a buffer gas (air, nitrogen, argon, helium, carbon dioxide, neon, propane, ethane, or SF_6 gas) was injected into the measuring cell to a certain pressure P_{BG}. Pressure P_{BG} in the measuring cell was measured with an elastic element pressure gage. The initial hydrogen concentration in the gases listed was preliminary measured with a chromatograph and was less than 1 ppm. When measuring the nonresonance background radiation as a function of buffer gas pressure P_{BG}, the signal $(I_a/I_p^2 I_s)^{1/2}$ being measured was calibrated against the signal from air. To this end, the buffer gas was pumped out of the cell after the measurements and air was supplied to the cell at room temperature and atmospheric pressure. Then, other conditions being the same, the nonresonance background from the air was measured to perform the calibration.

To study the interference between coherent scattering contributions from buffer gas molecules and molecules of the impurity, molecular hydrogen was injected in portions using a chromatographic syringe into the measuring cell containing the buffer gas under a certain pressure. Once a portion of the molecular gas had been introduced, the measurements were taken after no less than 15 min for the buffer gas and the hydrogen injected to uniformly mix together in the cell. The experiments were performed at room temperature.

3. RESULTS AND DISCUSSION

First, we studied nonresonance signal $[I_a/I_p^2 I_s]^{1/2}$ as a function of buffer gas pressure P_{BG}. The results are shown in Fig. 3. For helium, air, argon, carbon dioxide, SF_6 gas, and nitrogen, the $[I_a/I_p^2 I_s]^{1/2}$ versus P_{BG} curves are nearly linear but incline to the abscissa axis at different angles. The nonresonance background is the lowest for helium (Fig. 3, curve *8*) and neon (not shown in Fig. 3), while that for ethane and propane is many times higher. It should be noted that such a significant difference in nonresonance background basically makes it possible to study gas interdiffusion by the SRS–CARS technique, for example, the diffusion of ethane, propane, and SF_6 gas in helium and neon. In Fig. 3, the $[I_a/I_p^2 I_s]^{1/2}$ versus P_{BG} curves for ethane and propane are essentially nonlinear. This may be related to the nonlinear pressure dependence of the particle density at a

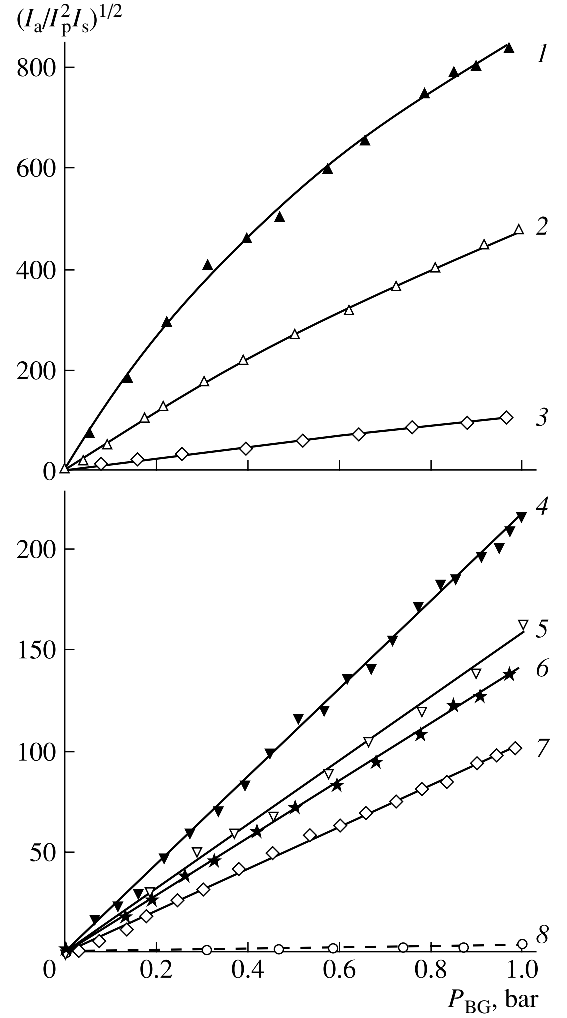


Fig. 3. Nonresonance signal $[I_a/I_p^2I_s]^{1/2}$ versus buffer gas pressure P_{BG} for (*1*) propane, (*2*) ethane, (*3*) nitrogen, (*4*) SF_6 gas, (*5*) carbon dioxide, (*6*) argon, (*7*) air, and (*8*) helium.

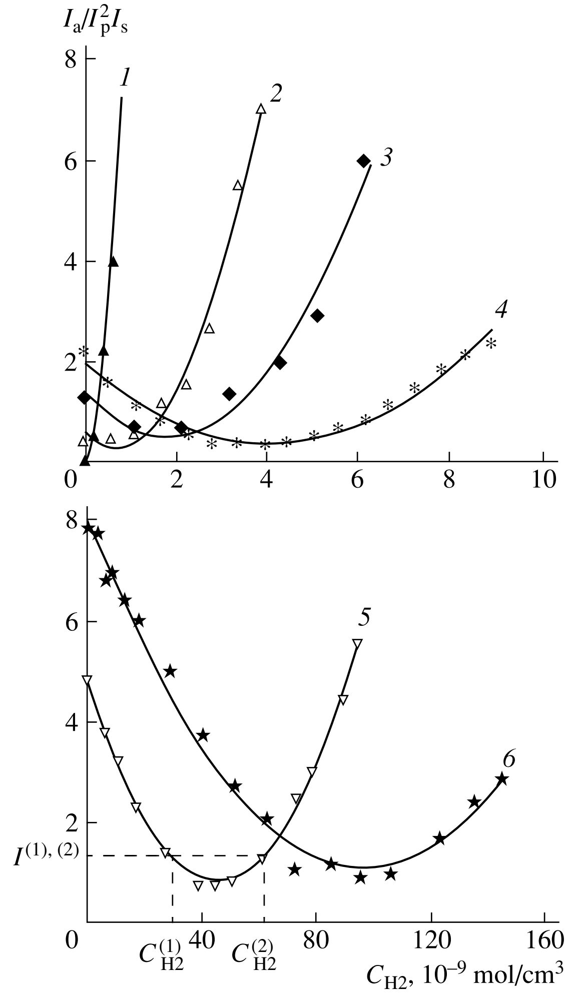


Fig. 4. SRS–CARS signal $(I_a/I_p^2I_s)$ versus molar concentration C_{H2} of hydrogen at an ethane pressure of (*1*) 0.12, (*2*) 0.20, (*3*) 0.30, (*4*) 0.52, (*5*) 0.76, and (*6*) 1.0 bar.

given temperature in these gases. The experimental data presented in Fig. 3 may be used to calculate the ratios of the nonresonance hyperpolarizabilities of particles in the gases studied. The results are summarized in the table.

In the experiments on interference between coherently scattered signals, the dependence of $I = I_a/I_p^2I_s$ on the hydrogen impurity concentration at various pressures of the buffer gases (nitrogen, argon, helium, air, carbon dioxide, neon, propane, ethane, and SF_6 gas) was studied. As an example, Figs. 4 and 5 plot $I_a/I_p^2I_s$ against the molar concentration C_{H2} of hydrogen with ethane (Fig. 4) and argon (Fig. 5) pressure taken as a parameter. At the ethane pressure $P_{eth} = 0.12$ bar, $I_a/I_p^2I_s$ is a monotonically increasing function of C_{H2} (Fig. 4, curve *1*). However, at higher ethane pressures in the measuring cell (P_{eth} = 0.2, 0.3, 0.52, 0.76, and 1 bar), the introduction of hydrogen in small amounts decreases the SRS–CARS signal. At a certain value of C_{H2}, C_{H2}^{min}, depending on P_{eth}, the scattered signal becomes minimal and subsequently grows with C_{H2} by a parabolic law. Thus, for ethane, the scattered signal is an ambiguous function of the hydrogen concentration even at a pressure of 0.2 bar (and above): at small levels of signal $I_a/I_p^2I_s$, the same signal corresponds to two hydrogen concentrations. In fact, in Fig. 4 (curve *5*), two hydrogen concentrations in ethane, $C_{H2}^{(1)}$ and $C_{H2}^{(2)}$, produce the same signal, $I^{(1),(2)}$. It also follows from Fig. 4 that the higher P_{eth}, the higher the hydrogen concentration that minimizes $I_a/I_p^2I_s$ as a function of C_{H2}. Similar results were also obtained for the mixtures of hydrogen

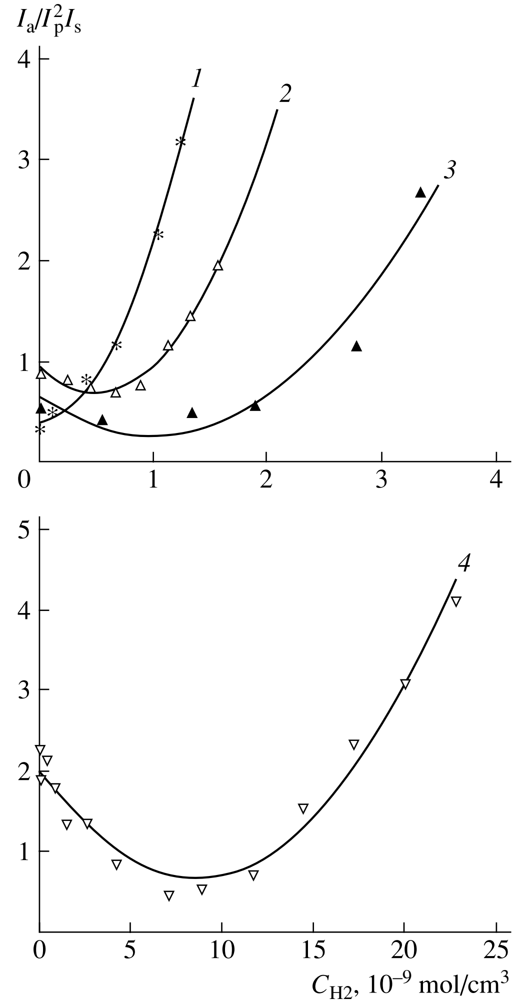


Fig. 5. SRS–CARS signal ($I_a/I_p^2I_s$) versus molar concentration C_{H2} of the hydrogen impurity at an argon pressure of (*1*) 0.48, (*2*) 0.60, (*3*) 0.76, and (*4*) 1.0 bar.

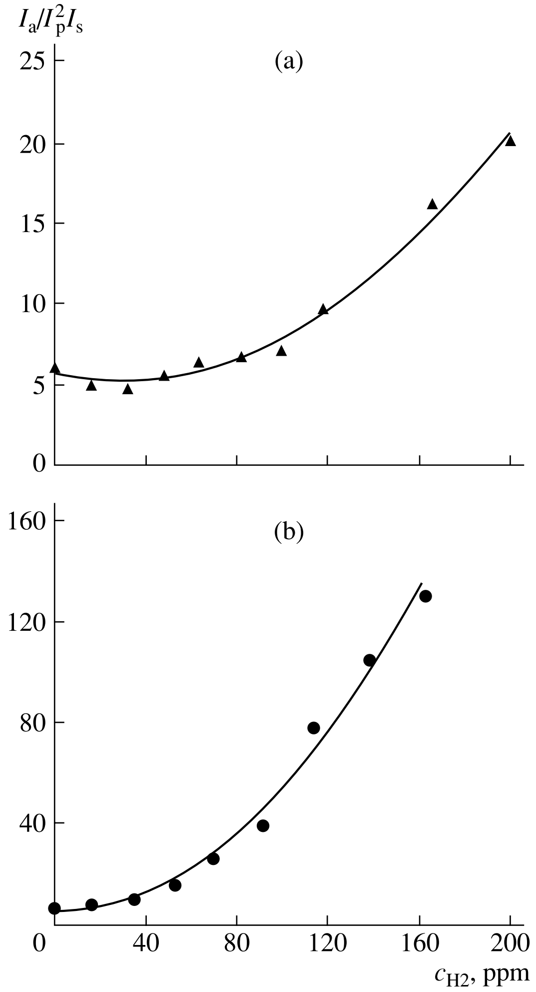


Fig. 6. SRS–CARS signal ($I_a/I_p^2I_s$) versus relative concentration c_{H2} of hydrogen in air under atmospheric pressure. The curves were obtained with the SRS oscillator cell filled with (a) pure H_2 and (b) H_2 (84%)–Ar (16%) gas mixture at a pressure of 4 bar (T = 295 K).

with argon, nitrogen, air, carbon dioxide, propane, and SF_6 gas. For all these mixtures, $I_a/I_p^2I_s$ monotonically increases with C_{H2} at pressures below 0.1 bar. However, the critical pressure above which the signal versus hydrogen concentration dependence becomes ambiguous depends on the buffer gas. In particular, for ethane, the critical pressure equals 0.2 bar (Fig. 4, curve *2*), while for argon, $I_a/I_p^2I_s$ is a monotonically increasing function of C_{H2} at pressures below 0.5 bar (Fig. 5, curve *1*). Remarkably, for the H_2–helium and H_2–neon mixtures, the experimental curves $I_a/I_p^2I_s(C_{H2})$ monotonically increase throughout the range of buffer gas pressures (from 0 to 1 bar).

These experimental results agree with formula (4) if the fact that parameter Δ depends both on the hydrogen pressure in the SRS oscillator and on the pressure of the gas mixture is taken into account. Indeed, considering that the pressure in the gas mixture is essentially the pressure of a buffer gas and using the results from [20], we can express Δ, in a first approximation, as

$$\Delta = \frac{a_{H2-H2}\rho_{H2-H2} - a_{H2-BG}\rho_{H2-BG}}{\Gamma}, \qquad (10)$$

where a_{H2-H2} is the coefficient that characterizes the sift of the Ω_{H2-H2} energy level due to collisions between hydrogen molecules ($a_{H2-H2} < 0$), ρ_{H2-H2} is the hydrogen density in the SRS oscillator, a_{H2-BG} is the coefficient characterizing the shift of the Ω_{H2-BG} energy level due to collisions between hydrogen molecules and buffer gas particles, and ρ_{H2-BG} is the density of the hydrogen–buffer gas mixture.

According to [20–22], coefficient a_{H2-BG} for the H_2–Ar and H_2–N_2 mixtures is negative; for the H_2–He and

Buffer gas-to-nitrogen nonresonance hyperpolarizability ratio γ_{BG}/γ_{N2} at λ_p = 532 nm (obtained in this work) and λ_p = 694.3 nm (calculated from the data in [16])

Buffer gas	γ_{BG}/γ_{N2}, λ_p = 532 nm	γ_{BG}/γ_{N2}, λ_p = 694.3 nm [16]
N_2	1	1
He	0.06	0.064
Ne	0.07	–
Air	0.98	–
Ar	1.3	1.14
CO_2	1.5	1.34
SF_6	2.0	1.84
NH_3	4.7	–
C_2H_6	5.3	4.6
C_3H_8	12.0	–

H_2–Ne mixtures, it is positive. Consequently, for the H_2–Ar and H_2–N_2 mixtures, Δ is negative at small ρ_{H2-BG} and positive when ρ_{H2-BG} is large. According to (4), this means that the experimental function $I_a/I_p^2 I_s(C_{H2})$ increases monotonically with hydrogen impurity concentration when buffer gas density ρ_{BG} is low (Figs. 4 and 5, curve *1*). At higher concentrations ρ_{BG} that meet the condition $\rho_{BG} \geq \rho_{BG}^{crit}$, where ρ_{BG}^{crit} is a certain critical buffer gas density, this function behaves in a different manner. As C_{H2} rises, the measured signal $I_a/I_p^2 I_s$ first declines, takes a minimum value, and then increases monotonically (Fig. 4, curves *2*–*6*; Fig. 5, curves *2*–*4*).

For the H_2–He and H_2–Ne mixtures, coefficient a_{H2-BG} is positive. Consequently, parameter Δ for these mixtures is negative at all ρ_{BG}. Then, formula (4) implies that the signal $I_a/I_p^2 I_s$ increases monotonically with the concentration of hydrogen in helium or neon at any pressure of the mixture. This speculation is confirmed by the experimental results obtained.

Critical buffer gas density ρ_{BG}^{crit}, above which $I_a/I_p^2 I_s$ as a function of the hydrogen concentration shows a minimum, may be raised either by increasing the hydrogen pressure in the SRS cell or by adding a buffer gas that decreases the frequency of the Q_{01} (1) vibrational transition (e.g., argon). The dependence $I_a/I_p^2 I_s(c_{H2})$ (where $c_{H2} = n_{H2}/N$ is the relative hydrogen concentration and $N = 2.68 \times 10^{19}$ particles/cm^3 is the Loschmidt number) experimentally found for air under atmospheric pressure with the help of the SRS oscillator at P_{H2} = 4 bar (Fig. 6a) exhibits a minimum. The same dependence obtained with the cell of the SRS converter (T = 295 K) at a total pressure of the H_2 (84%)–Ar (16%) mixture of 4 bar monotonically grows (Fig. 6b). Thus, the concentration of hydrogen in dense gas mixtures can be uniquely determined by appropriately choosing the composition and pressure of a hydrogen–buffer gas mixture in the SRS cell. The application of the computerized laser system for SRS–CARS diagnostics [7] and optimization of the gas composition and pressure in the cell of the SRS oscillator allowed us to achieve the sensitivity of hydrogen detection in air at atmospheric pressure as high as 5 ppm. This value is more than one order of magnitude better than that obtained in [10].

CONCLUSIONS

In this work, we report the results for gas mixture diagnostics by the SRS–CARS technique. The dependence of the nonresonance SRS–CARS signal on the pressure of various gases is studied. Interference between the resonance and nonresonance nonlinear susceptibilities is shown to significantly affect the signal recorded when low hydrogen concentrations in dense gas mixtures are monitored by the SRS–CARS technique. Because of this effect, in gas mixtures (such as H_2–Ar, H_2–N_2, H_2–C_2H_6, and others) where the frequency of the Raman vibrational transition in H_2 decreases with increasing gas mixture pressure, the scattered signal intensity versus hydrogen concentration becomes ambiguous after a certain pressure has been reached. In H_2–He and H_2–Ne mixtures, where the frequency of this vibrational transition increases with pressure, the ambiguity is absent. It is shown that the ambiguity in the SRS–CARS signal can be removed through a frequency offset of the biharmonic laser pumping and by appropriately choosing the pressure and composition of the compressed gas in the SRS oscillator.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, grant nos. 04-02-96011 and 01-02-96461.

REFERENCES

1. K. Okuno, T. Uda, S. O’Hira, and Y. Naruse, J. Nucl. Sci. Technol. **28**, 509 (1991).
2. K. Kalli, A. Othonos, and C. Christofides, Rev. Sci. Instrum. **68**, 3544 (1997).
3. Georg. M. Mikheev and G. M. Mikheev, Zavod. Lab. Diagnostika Mater. **67** (6), 38 (2001).
4. G. M. Mikheev, Georg. M. Mikheev, A. N. Kulikov, and S. F. Nikitin, Proc. SPIE **4749**, 178 (2002).
5. G. M. Mikheev, Georg. M. Mikheev, E. G. Fateev, and A. Yu. Popov, Zh. Tekh. Fiz. **72** (10), 73 (2002) [Tech. Phys. **47**, 1277 (2002)].
6. E. I. Voronina, V. E. Privalov, and V. G. Shemanin, Pis’ma Zh. Tekh. Fiz. **30** (5), 14 (2004) [Tech. Phys. Lett. **30**, 178 (2004)].

7. G. M. Mikheev, T. N. Mogileva, A. Yu. Popov, and D. G. Kaluzhnyĭ, Prib. Tekh. Éksp., No. 2, 101 (2003).
8. G. M. Mikheev and T. N. Mogileva, Int. J. Hydrogen Energy **24**, 833 (1999).
9. P. R. Regnier and J. P. E. Taran, Appl. Phys. Lett. **23**, 240 (1973).
10. A. A. Ivanov, G. A. Polyakov, and V. B. Voronin, Izv. Ross. Akad. Nauk, Ser. Fiz. **57**, 165 (1993).
11. G. M. Mikheev and T. N. Mogileva, Kvantovaya Élektron. (Moscow) **23**, 943 (1996).
12. G. M. Mikheev, Georg. M. Mikheev, T. N. Mogileva, and D. G. Kaluzhnyĭ, Kvantovaya Élektron. (Moscow) **32**, 39 (2002).
13. S. N. Bagaev, A. K. Dmitriyev, S. V. Chepurov, *et al.*, Laser Phys. **11**, 1270 (2001).
14. F. Benabid, J. C. Knight, G. Antonopolos, and P. S. J. Russell, Science **298**, 399 (2002).
15. S. O. Konorov, A. B. Fedotov, D. A. Sidorov-Biryukov, *et al.*, J. Raman Spectrosc. **34**, 688 (2003).
16. S. A. Akhmanov and N. I. Koroteev, *Methods of Nonlinear Optics in Light Scattering Spectroscopy* (Nauka, Moscow, 1981) [in Russian].
17. R. J. Hall and A. C. Eckbreth, Laser Appl. **5**, 213 (1984).
18. G. M. Mikheev, D. G. Kaluzhnyi, and A. Yu. Popov, Pis’ma Zh. Tekh. Fiz. **29** (24), 62 (2003) [Tech. Phys. Lett. **29**, 1040 (2003)].
19. G. M. Mikheev, D. I. Maleev, and T. N. Mogileva, Kvantovaya Élektron. (Moscow) **19**, 45 (1992).
20. A. D. May, V. Degen, J. C. Stryland, and H. L. Welsh, Can. J. Phys. **39**, 1769 (1961).
21. J. Ph. Berger, R. Saint-Loup, H. Berger, *et al.*, Phys. Rev. A **49**, 3396 (1994).
22. P. M. Sinclair, J. Berger, X. Michaut, *et al.*, Phys. Rev. A **54**, 402 (1996).

Translated by A. Khzmalyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 227–231. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 89–93.
Original Russian Text Copyright © 2005 by Petrova, Chichinadze, Shaverdova.

OPTICS, QUANTUM ELECTRONICS

Kinetics of the Weigert Effect in Azo Dyes Embedded in Polymeric Matrices with Different Activities

S. S. Petrova, N. M. Chichinadze, and V. G. Shaverdova
Institute of Cybernetics, Georgian Technical University, Academy of Sciences of Georgia,
Tbilisi, 380086 Georgia

Received April 6, 2004

Abstract—Experimental data for the photoanisotropy kinetics in azo dyes embedded in various polymer matrices are reported. The Weigert effect in these dyes is shown to depend on the polarized actinic radiation wavelength and matrix constitution. The effect of dark relaxation in the dyes is found. *© 2005 Pleiades Publishing, Inc.*

Light-sensitive media that become anisotropic when exposed to polarized light find application in polarization holography [1]. An example of such anisotropic media is organic azo dyes embedded in polymeric matrices. Previous studies [2, 3] showed that the photoanisotropic properties of these media depend on the constitution of the matrix and dye. Under normal conditions, most azo dyes exist as a combination of *trans*- and *cis*-isomers, the former prevailing because of their greater volume. When absorbed, the energy of visible or near-UV photons is insufficient for N–N bond breaking in a dye molecule; however, the absorbed energy loosens the bond to the point where the end groups may rotate around it, giving rise to coordination *trans*- or *cis*-isomerism [4]. Further irradiation raises the concentration of *cis*-isomers and decreases that of *trans*-isomers. Once the irradiation has been terminated, *cis*-isomers spontaneously turn into energetically more favorable *trans*-isomers. Although the quantum yield of *cis*- and *trans*-isomers in solid matrices is two to three times higher than in liquid solutions, the efficiency of the reverse transition does not depend on the aggregative state of the medium and recovery to the initial state may be difficult under certain conditions [5].

Irradiation by linearly polarized light considerably increases the probability that those randomly oriented dye molecules the oscillation axes of which make a small angle with the oscillation axis of the electric vector of the light will take part in photochemical reactions. According to the phenomenological model, the number of elementary cells oriented within $d\theta$ near θ is

$$dN = \frac{N}{\pi}d\theta.$$

After the action of the polarized light, the number of the cells remaining in the initial state will be

$$dN'(\theta) = e^{-\rho(H_1 + H_2)}dN(\theta)$$

on the assumption that a photochemical reaction obeys the exponential law. Here, N is the total number of the cells and ρ is the photochemical reaction efficiency. Then, the number of cells of the same orientation that have reacted with the light is given by

$$dN''(\theta) = dN(\theta) - dN'(\theta).$$

During irradiation by linearly polarized light, a molecule enters into a photochemical reaction and changes orientation until the orientation of its absorbing oscillator becomes nearly orthogonal to the polarization of the incident radiation [6].

The photoinduced anisotropy was measured in real time with a photometric setup that writes anisotropy data at the wavelength $\lambda = 4416$ Å and reads them out at $\lambda = 6328$ Å in one measurement cycle. The measurand in this work was effective anisotropy A_{eff}, which equals the magnitude of the anisotropic invariant of the Jones matrix:

$$A_{\text{eff}} = |\gamma^2|$$
$$= \frac{1}{2}\exp\{-2\kappa d(\overline{n\tau})\}[\cosh\kappa d\Delta(n\tau) - \cos\kappa d\Delta n],$$

where $\kappa d(\overline{n\tau})$, $\kappa d\Delta(n\tau)$, and $\kappa d\Delta n$ are, respectively, the average absorption, dichroism, and birefringence of the medium.

Anisotropic invariant $|\gamma^2|$ is numerically equal to the transmission of the sample placed between two crossed polarizers in the case when the induced anisotropy axis and the axis of one polarizer make an angle of 45°. From the expression for A_{eff}, it follows that this parameter takes into account the net contribution of dichroism, birefringence, and scalar absorption to the induced anisotropy [7].

Based on experimental data, we constructed kinetic curves $A_{\text{eff}} = f(t)$ and calculated the photoanisotropic sensitivity. This sensitivity is defined as the reciprocal

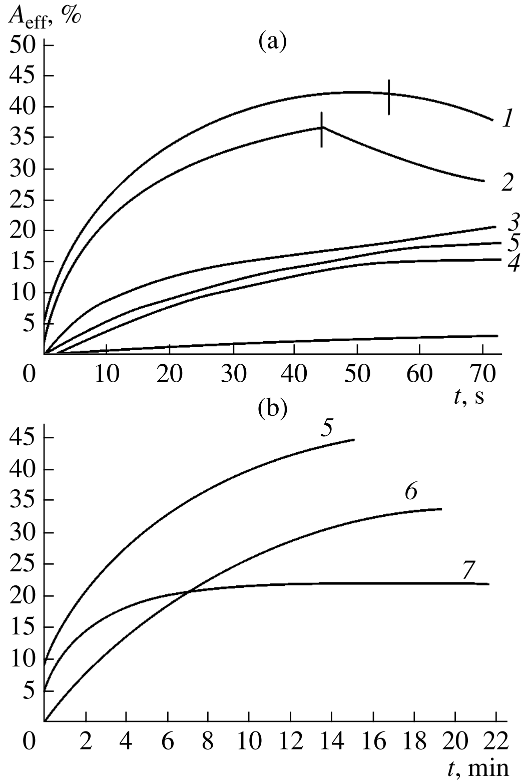


Fig. 1. Photoanisotropy kinetics $A_{eff} = f(t)$ in the gelatin matrix for dyes (*1*) chrome yellow "K," (*2*) chrome yellow "Z," (*3*) MPY, (*4*) MPY "M," (*5*) AYF, (*6*) chrysophenine, and (*7*) bright yellow.

of the exposure that is necessary for A_{eff} to exceed an initial level Q_0 by a certain amount Q_{cr}, the so-called photoanisotropic sensitivity criterion:

$$S = \{H_Q = Q_0 + Q_{cr}\}^{-1}, \quad Q = \log A_{eff}.$$

It is usually assumed that $Q_0 = 0.3$, which reflects the sensitivity threshold of most of the devices, and $Q_{cr} = 0.2$. The photoanisotropic sensitivity characterized in such a way makes it possible to contrast various photoanisotropic materials (the same dye embedded in different matrices and the same matrix incorporating different dyes). The dark relaxation was estimated from parameter $\tau_{1/2}$, which is the time period over which A_{eff} is halved.

In this work, we studied polymeric matrices with different activities (gelatin, polyacrylic acid (PA), a PA–caprolactam (CL) polymeric complex, nitrocellulose (NC), an epoxy polymer (EP), polyvinyl pyrrolidone (PVP), polyvinyl alcohol (PVA), polystyrene (PS)) and 14 azo dyes. The concentrations of the dyes and polymers were 0.3 and 3%, respectively. The homogeneous solutions of the dyes and polymers were mixed together and applied on glass substrates. The thickness of the films dried was ≈10 μm.

Consider the photoanisotropic properties of each of the matrix–dye systems in detail. The gelatin matrix is neutral. It has the form of a globule with a variable space between individual fragments of a gelatin macromolecule. The water-soluble dyes considered in this work dissociate in an aqueous solution and may either add to amphoteric gelatin molecules or diffuse in the free space between individual fragments of a gelatin molecule [8].

A specific feature of photoanisotropy in this matrix is that the accumulation of the *cis*-forms proceeds gradually; that is, it takes a considerable time for anisotropy to reach a maximum. The reverse process, *cis*-to-*trans* dark relaxation, proceeds very slowly, since the material properties are stable. In the samples covered by a water-proof layer, the induced anisotropy persists for a long time. Therefore, this material is used in polarization holography, in production of polarization optics and anisotropic diffraction gratings, etc. It seems likely that the stability of the material is associated with a decrease in the free volume of the rigid matrix, which hinders molecule reversal. The fact that the induced anisotropy disappears upon moistening of the sample favors this supposition. The material is reversible, i.e., allows for multiple data write/read. The photoanisotropy may be erased by exposing the samples to radiation of the same wavelength but orthogonal polarization (Figs. 1a, 1b).

Figure 1 shows that mordant monoazo dyes chrome yellow "K" and chrome yellow "Z" (curves *1*, *2*) exhibit a rapid rise in the effective anisotropy (relatively high values of A_{eff}) even early in the illumination. After the illumination, the curves asymptotically fall.

In mordant pure yellow (MPY) diazo dye (curve *3*), as well as in its methyl analogue MPY "M" (curve *4*), the number of possible structural isomers grows, providing a high absolute value of A_{eff}. The same fact, in our opinion, decreases the light sensitivity of the material and causes a rapid increase in the anisotropy with exposure. The dark relaxation proceeds very slowly.

The light sensitivity of the diazo dyes acid yellow for fulling (AYF) (curve *5*), chrysophenine (curve *6*), and bright yellow (curve *7*) is low, and so is the absolute value of the induced anisotropy early in the illumination. However, earlier studies [9] indicate that high exposures result in a drastic increase in the effective anisotropy. This to a greater extent refers to bright yellow dye, where a hydroxyl group is bonded to an azo group, making the formation of hydrogen bridges possible. Because of these competing factors, the anisotropy at low exposures is low. At high exposures, the effect of *trans–cis* isomerism becomes appreciable. Therefore, the photoanisotropy greatly increases in chrysophenine and AYF (Fig. 1b), since the electron donation power of a hydroxyl group is much higher than that of an ethoxy group.

PVA and PS are neutral hydrogen-containing matrices. In both, small dye molecules interact with segments of polymer macromolecules, presumably via van der Waals forces and diffusion of the small molecules into the free volume in the macromolecules (Figs. 2, 3).

For all the dyes embedded in these matrices, A_{eff} rapidly grows early in the illumination. The sensitivity of these materials reaches 50 J/m, and the dark relaxation takes from 4.5 to 20 s.

PVP, NC, EP, PA, and the PA–CL complex are chemically active matrices. The PVP matrix with dyes tropeolin 00, metanil yellow, chrome yellow “K,” chrome yellow “Z,” MPY, and MPY “M” offers a high sensitivity and high absolute values of effective anisotropy A_{eff}. Here, the quick rise in the anisotropy may be attributed to the fact that, owing to the nitrogen atom present in a pyrrole ring of the polymeric molecule, it may form complexes with the dye molecules and isomerization due to polarized photons is facilitated [10].

Figure 4 shows the kinetic curves of the effective anisotropy for the dye–PVP systems. For diazo dyes, such as MPY and MPY “M,” embedded in PVP, the run of the curves is the same as for these dyes in the gelatin matrix: the absolute value of A_{eff} grows with exposure and becomes considerable within 15–20 min after the beginning of illumination.

Let us turn to NC, which also is a chemically active matrix. Its chemical activity is apparently related to the presence of nitrogen atoms and shows up in a change in the absorption spectra of the dyes compared with their absorption spectra in the inactive matrices.

Figure 5 demonstrates the kinetic curves of the effective anisotropy in NC. The run of the curves is nearly identical for all the dyes, but curve *1* for dimethyl yellow dye increases very steeply at low exposures. The light sensitivity of the material is fairly high, 50 J/m. Further illumination increases the photoanisotropy only slightly, and saturation is observed at relatively low absolute values of A_{eff}. The time of dark relaxation is roughly the same, ≈13 s.

The epoxy polymer is a polymeric phenolic ether with epoxy end groups, which are highly reactive and readily enter into a chemical reaction, breaking the bonds.

As follows from Fig. 6, the photoanisotropy rapidly increases early in the illumination and, consequently, the sensitivity of this dye–matrix system is high, ≈100 J/m. The absolute value of the photoanisotropy is also appreciable, ≈37%. The dark relaxation lasts ≈1 min, indicating that the molecules of the dye in the modified state are stable.

The next chemically active matrix in which all the water-soluble dyes were embedded was the specially synthesized PA–CL complex (for the synthesis procedure, see [11]). The interaction of the dyes with this matrix depends on their structures. The absorption spectra of acid monoazo dyes (metanil yellow, tropeolin 00, and methyl orange) exhibit a bathochromic shift because of their interaction with the matrix. The spectra of the dyes brilliant yellow, chrysophenine, and AYF are modified in a more complicated way. The mordant dyes in the CL matrix are somewhat bleached, which also follows from their absorption spectra. The spectra of the dyes Indian yellow and azoflavin N remain unchanged.

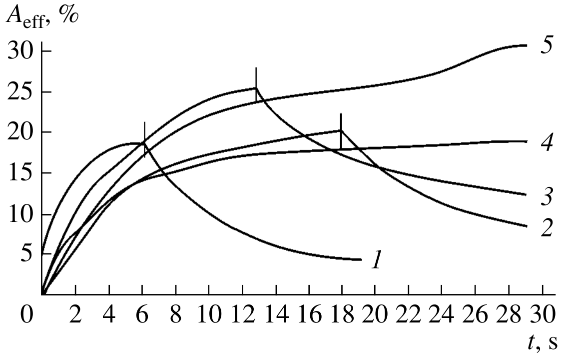


Fig. 2. Photoanisotropy kinetics $A_{\text{eff}} = f(t)$ in the PVA matrix for dyes (*1*) methyl orange, (*2*) tropeolin 00, (*3*) methyl yellow, (*4*) MPY “M,” and (*5*) alizarin yellow.

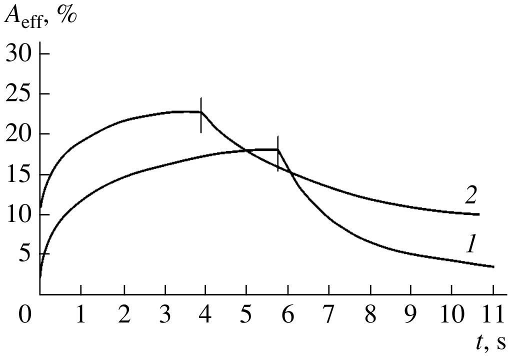


Fig. 3. Photoanisotropy kinetics $A_{\text{eff}} = f(t)$ in the PC matrix for dyes (*1*) dimethyl yellow and (*2*) methyl red.

The irradiation of these samples by linearly polarized light induces a noticeable anisotropy, although the light sensitivity of these dyes in the CL matrix is moderate, indicating that isomerization competes with a chemical reaction in these systems (Fig. 7). The time of dark relaxation has a significant spread: from 1 s for chrome yellow “K” (curve *5*) to 40 s for AYF (curve *6*).

The last of the matrices studied was the PA matrix. In its chemical properties, PA is akin to polybasic saturated acids. The interaction of the dyes with this matrix is similar to the interaction with the CL matrix but is more pronounced. Because of the acidity of this matrix, molecules of dyes tropeolin 00, metanil yellow, and methyl orange are protonated with the formation of a quinoid benzene ring. The bathochromic shift here is more pronounced. Similarly, the bleach of the mordant azo dyes is more intense. The absorption spectra of the

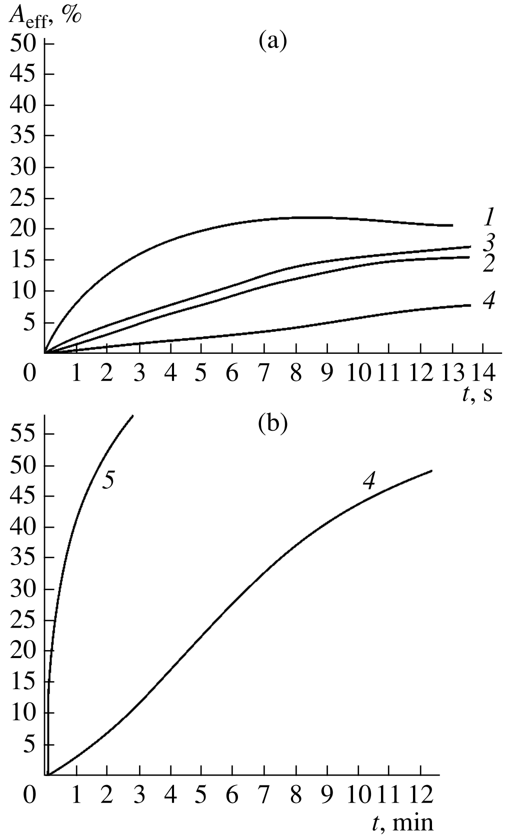


Fig. 4. Photoanisotropy kinetics $A_{eff} = f(t)$ in the PVP matrix for dyes (*1*) tropeolin 00, (*2*) chrome yellow "K," (*3*) metanil yellow, (*4*) MPY "M," and (*5*) MPY.

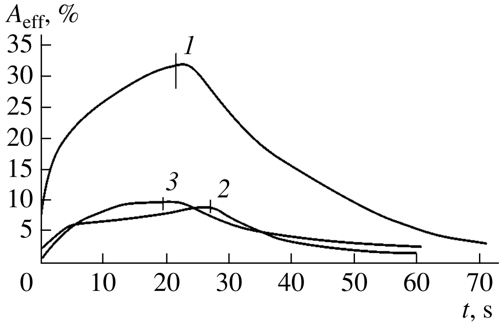


Fig. 5. Photoanisotropy kinetics $A_{eff} = f(t)$ in the nitrocellulose matrix for dyes (*1*) dimethyl yellow, (*2*) methyl red, and (*3*) benzyl orange.

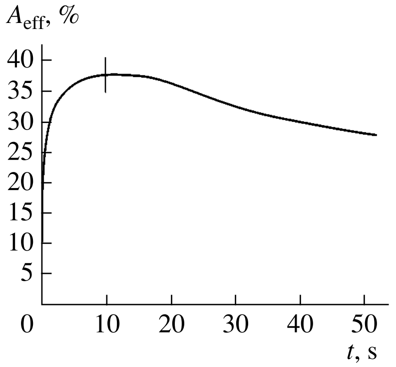


Fig. 6. Photoanisotropy kinetics $A_{eff} = f(t)$ in the epoxy polymer matrix for dimethyl yellow dye.

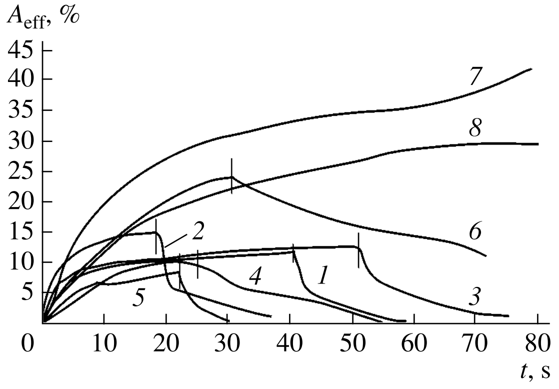


Fig. 7. Photoanisotropy kinetics $A_{eff} = f(t)$ in the PA–CL matrix for dyes (*1*) tropeolin 00, (*2*) metanil yellow, (*3*) Indian yellow, (*4*) azoflavin N, (*5*) chrome yellow "K," (*6*) AYF, (*7*) bright yellow, and (*8*) MPY "M." The vertical bars indicate the beginning of relaxation.

dyes brilliant yellow and chrysophenine are irregular. The samples exhibit a wide variety of colors.

The dyes whose molecules incorporate strongly accepting groups, such as NO_2 and NO nitrogroups, have nearly the same color in both the initial and protonated state. These are Indian yellow and azoflavin N.

The irradiation of samples with all the dyes considered by actinic radiation from a He–Cd laser (λ = 441.6 nm) does not induce anisotropy, irrespective of whether their absorption spectra are modified or not.

Thus, we can state with assurance that the chemical reaction with the PA matrix makes the photochemical reaction of *cis–trans* isomerization impossible in all the dyes studied in this work. In these dyes, this reaction is a basic photochemical reaction responsible for photoanisotropy [12].

Our results suggest that the matrix is of crucial importance in development of a photoanisotropic material.

All the matrices used in this work can be subdivided into those entering into a chemical reaction with dye molecules and those that are chemically inactive.

The latter form multicomponent complexes with dye molecules. Sometimes, the resulting effect is beneficial: the photoanisotropic activity and light sensitivity of the dye–polymer system are improved. Such matrices are PVP, NC, and the epoxy polymer.

It seems that complexation does not modify the molecular constitution of the dyes in these systems and irradiation by linearly polarized actinic light favors photoisomerization and, accordingly, imparts to them photoanisotropic properties.

In a number of cases, chemical reactions in the dye–polymer system change the π-electron constitution of dye molecules with the formation of the resonant quinoid structure.

If isomerization competes with the formation of the quinoid–hydrazone form of a dye, actinic radiation will induce anisotropy but its absolute value, as well as the light sensitivity of the material, will be low. The PA–CL polymer is an example of such matrices.

If isomerization is completely suppressed, as in the PA matrix, polarized light does not cause anisotropy.

In the neutral matrices, such as gelatin, PVA, and PS, the free volume between fragments of a polymer macromolecule where small dye molecules may diffuse is of particular importance. The photoanisotropic properties of such materials will depend on their isomerization facility, i.e., on the rigidity of the matrix, temperature, moisture content, etc.

Research in this area is being continued. Related results will be reported in following publications.

REFERENCES

1. Sh. D. Kakichashvili, *Polarization Holography* (Nauka, Leningrad, 1989), p. 6 [in Russian].
2. A. Ya. Zheltov, B. I. Stepanov, and V. G. Shaverdova, Zh. Prikl. Spektrosk. **52**, 280 (1990).
3. G. A. Kakuridze, V. G. Shaverdova, Ya. A. Shvaĭtser, and I. D. Shatalin, Pis'ma Zh. Tekh. Fiz. **16** (24), 59 (1990) [Sov. Tech. Phys. Lett. **16**, 948 (1990)].
4. A. N. Terenin, *Photonics of Dye Molecules* (Nauka, Leningrad, 1967), pp. 208–209 [in Russian].
5. J. Guillet, *Polymer Photophysics and Photochemistry: An Introduction to the Study of Photoprocesses in Macromolecules* (Cambridge Univ. Press, Cambridge, 1985; Mir, Moscow, 1988), p. 115.
6. Sh. D. Kakichashvili and Ya. A. Shvaĭtser, *Photoanisotropic and Photogyrotropic Phenomena in Condensed Media and Polarization Holography* (Tbilisi, 1988), pp. 28–29 [in Russian].
7. A. I. Balabanov, G. A. Kakauridze, Sh. D. Kakichashvili, and V. G. Shaverdova, Opt. Spektrosk. **67**, 409 (1989) [Opt. Spectrosc. **67**, 238 (1989)].
8. T. B. Brill, *Light: Its Interaction with Art and Antiquities* (Plenum, New York, 1980; Mir, Moscow, 1983).
9. V. G. Shaverdova, B. I. Stepanov, and A. Ya. Zheltov, *Vector and Three-Dimensional Holography* (Tbilisi, 1990), pp. 136–141 [in Russian].
10. J. Guillet, *Polymer Photophysics and Photochemistry: An Introduction to the Study of Photoprocesses in Macromolecules* (Cambridge Univ. Press, Cambridge, 1985; Mir, Moscow, 1988), p. 20.
11. N. M. Chichinadze, USSR Inventor's Certificate No. 1,179,923 (1963).
12. *Indicators,* Ed. by E. Bishop (Pergamon, Oxford, 1972; Mir, Moscow, 1976).

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 232–236. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 94–98.
Original Russian Text Copyright © 2005 by Golyatina, Tkachev, Yakovlenko.

OPTICS, QUANTUM ELECTRONICS

Analysis of a Heat Wave Induced by Laser Radiation Absorption in an Optical Fiber on the Basis of a 2D Nonstationary Heat Conduction Equation

R. I. Golyatina, A. N. Tkachev, and S. I. Yakovlenko

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

Received June 2, 2004

Abstract—A heat wave resulting from the absorption of laser radiation in the core of an optical fiber is studied using a nonstationary 2D heat conduction equation. The velocity of the wave as a function of the laser intensity is determined, and the threshold intensity generating the heat wave is calculated. At high intensities, the velocity of the wave can be qualitatively described by a well-known formula from combustion theory; i.e., the velocity is shown to be proportional to the square root of the radiation intensity. The analytical threshold laser intensities closely agree with the available experimental data. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

It has long been known that intense laser radiation may have a dramatic impact on the physical parameters of a transparent condensed medium [1]. In particular, when the laser intensity exceeds a certain threshold, the absorption factor increases steeply. Under the conditions of optical discharge in condensed media [2, 3] or gases [4], this gives rise to an absorption wave propagating toward the laser radiation.

In recent years, much interest has arisen in the same phenomenon in optical fibers [5–15]. If the absorption factor rises somewhere in a fiber, heating of this place further enhances the absorption. Due to the heat conduction process, such a heat absorption wave (HAW) propagates toward the laser radiation.

Previously, the HAW was considered in a simplified stationary 1D approximation using a coordinate system related to the wave [7, 8]. In this study, HAW analysis is based on a nonstationary 2D heat conduction equation. The dependences of the HAW velocity on the laser intensity are derived, and the threshold intensities at which the HAW arises are determined.

MATHEMATICAL MODEL

Heat conduction equation. A nonstationary 2D model of an HAW in the cylindrical coordinates (r, z) within the rectangular domain $0 \le r \le r_1$, $0 \le z \le l$ (r_1 and l are the outer radius and length of the fiber, respectively) is described by the heat conduction equation

$$c_p(T)\rho(T)\frac{\partial}{\partial t}T(t,z,r) = \frac{\partial}{\partial z}\left[k(T)\frac{\partial}{\partial z}(T(t,z,r))\right] + \frac{1}{r}\frac{\partial}{\partial r}\left[rk(T)\left(\frac{\partial}{\partial r}(T(t,z,r))\right)\right] + \alpha(T)I(t,z,r), \quad (1)$$

and the radiation transfer equation

$$\frac{\partial}{\partial z}I(z,r) = -\alpha(T)I(z,r). \quad (2)$$

Here, z is the coordinate along the fiber, r is the radial coordinate, $c_p(T)$ is the specific heat at constant pressure, I is the intensity (energy flux density) of the laser radiation, $\alpha(T)$ is the absorption factor, $k(T)$ is the thermal conductivity, and $\rho(T)$ is the density of the material. The time dependence of the intensity is embodied in the time dependence of the temperature. The temperature dependence of the absorption factor was chosen as follows:

$$\alpha(T) = \begin{cases} 0, & T < T_1 \\ \alpha_p(T-T_1)/(T_p-T_1), & T_1 \le T \le T_p \\ \alpha_p, & T > T_p, \end{cases}$$

where α_p is the peak value of the absorption factor, T_p is the temperature at which α reaches a peak value, and T_1 is the temperature at which the absorption starts rapidly growing. We assume that $(T_p - T_1) \ll T_p$, so that the results are virtually independent of T_1. According to [7, 8], we put $\alpha_p = 560\ \text{cm}^{-1}$ and $T_p = 2000°\text{C}$. If α_p is far from this value, the result of calculation become inconsistent with experimental data. We also take T_1 = 1700°C.

The dependence of the specific heat on temperature and on the parameters of the phase transition was represented in the form

$$c_p(T) = c_0(T) + \Delta c(T, T_\mathrm{m}, \Delta T_\mathrm{m}, \Delta H_\mathrm{m}) + \Delta c(T, T_\mathrm{p}, \Delta T_\mathrm{p}, \Delta H_\mathrm{p}).$$

Here, function $c_0(T)$ describes the temperature dependence of the specific heat in the absence of the phase transition,

$$\Delta c(T, T_0, \Delta T_0, \Delta H) = (\Delta H_0/\pi^{1/2}\Delta T_0)\exp\{-[(T - T_0)/\Delta H_0]^2\}$$

is the discontinuity of the specific heat at the point of the phase transition, T_m is the melting point, T_p is the temperature at which the absorption steeply increases, ΔH_m is the heat of melting, and ΔH_p is the heat of the phase transition when the absorption is enhanced. The value of ΔT_0 characterizes the width of the phase transition.

For glass, we used the following values of the parameters in our calculation [7, 8, 16]: $\rho(T)$ = 2.2 g/cm^3, k = 0.02 W/(cm K), $c_0(T)$ = 0.74 J/g K, T_m = 1600°C, T_p = 2000°C, ΔH_m = 142 J/g, ΔH_p = 142 J/g, ΔT_m = 100 K, and ΔT_p = 100 K.

Boundary and initial conditions. It is assumed that radiation of intensity I_0 enters the medium at $z = 0$; i.e., $I(t, 0, r) = I_0(r)$. Then, from Eq. (2), we have

$$I(t, z) = I_0\exp\left(-\int_0^z \alpha(T(t, z'))dz'\right).$$

However, this expression is inconvenient for numerical integration, since it contains the desired quantity $T(t, z)$. In the calculations, a step radial distribution of the input intensity was therefore used: $I_0(r) = I_0$ at $r < r_0$ and $I_0(r) = 0$ at $r \geq r_0$. Heat removal from the fiber surface was assumed to be absent,

$$\frac{\partial}{\partial r}T(t, z, r)\bigg|_{r = r_1} = 0, \quad \frac{\partial}{\partial z}T(t, z, r)\bigg|_{z = 0,\, z = l} = 0.$$

The initial conditions corresponded to the step distribution: $T(t, z, r)|_{t = 0} = T_0$ at $z < z_\mathrm{p}$ and $T(t, z, r)|_{t = 0} = T_\mathrm{p}$ at $z \geq z_\mathrm{p}$. Here, z_p is the coordinate of the point where the initial perturbation occurs. The temperature was set equal to T_0 = 20°C.

SOLUTION ALGORITHM FOR THE 2D PROBLEM

Let us construct a five-point finite-difference second-order approximant for Eq. (1) using a spatial grid uniform in the z direction with a step hz and a grid quasi-uniform in the r direction [16, 17],

$$\psi(\alpha, t) = r_1\frac{\ln^\alpha(1.01 + \alpha t) - \ln^\alpha(1.01)}{\ln^\alpha(1.01 + \alpha) - \ln^\alpha(1.01)},$$

$$\text{where} \quad t = \frac{i - 1}{Nr - 1}.$$

Then, for an internal grid node (i, j) $(1 < i < Nr,\ 1 < j < Nz)$, we have

$$c_p(T_{i,j})\rho(T_{i,j})\frac{\partial}{\partial t}T_{i,j} = P^1_{i,j}T_{i,j-1} + P^3_{i,j}T_{i,j+1} + P^2_{i,j}T_{i-1,j} + P^4_{i,j}T_{i+1,j} - (Pr^5_{i,j} + Pz^5_{i,j})T_{i,j} + \alpha(T_{i,j})I_{i,j},$$

where

$$P^1_{i,j} = \frac{K_{i,j-\frac{1}{2}}}{h_z^2}, \quad P^3_{i,j} = \frac{K_{i,j+\frac{1}{2}}}{h_z^2}, \quad Pz^5_{i,j} = P^1_{i,j} + P^3_{i,j},$$

$$P^2_{i,j} = \frac{r_{i-\frac{1}{2}}K_{i-\frac{1}{2},j}}{hr_{i-1}\left(r_{i+\frac{1}{2}} - r_{i-\frac{1}{2}}\right)r_i}, \quad P^4_{i,j} = \frac{r_{i+\frac{1}{2}}K_{i+\frac{1}{2},j}}{hr_i\left(r_{i+\frac{1}{2}} - r_{i-\frac{1}{2}}\right)r_i},$$

$$Pr^5_{i,j} = P^2_{i,j} + P^4_{i,j},$$

and $hr_i = r_{i+1} - r_i$, and $r_{i+\frac{1}{2}}$ is the midpoint of the quasi-uniform grid interval, which is found via the same transformation ψ that specifies the grid itself:

$$r_{i+\frac{1}{2}} = \psi\left(\frac{i - \frac{1}{2}}{Nr - 1}\right).$$

Taking into consideration that, at $r = 0$, Eq. (1) appears as

$$c_p(T)\rho(T)\frac{\partial}{\partial t}T(t, z, r) = \frac{\partial}{\partial z}\left[k(T)\frac{\partial}{\partial z}(T(t, z, r))\right] + 2k(T)\frac{\partial^2}{\partial r^2}(T(t, z, r)) + \alpha(T)I(t, z, r)$$

and that the Neumann boundary conditions are set on the boundaries, we obtain the difference scheme coefficients at the corresponding boundaries in the following form:

for $i = 1, \ldots, Nr$,

$$P^1_{i,1} = 0, \quad P^3_{i,1} = \frac{2K_{i,\frac{1}{2}}}{h_z^2}, \quad P^1_{i,Nz} = \frac{2K_{i,Nz-\frac{1}{2}}}{h_z^2},$$

$$P^3_{i,Nz} = 0;$$

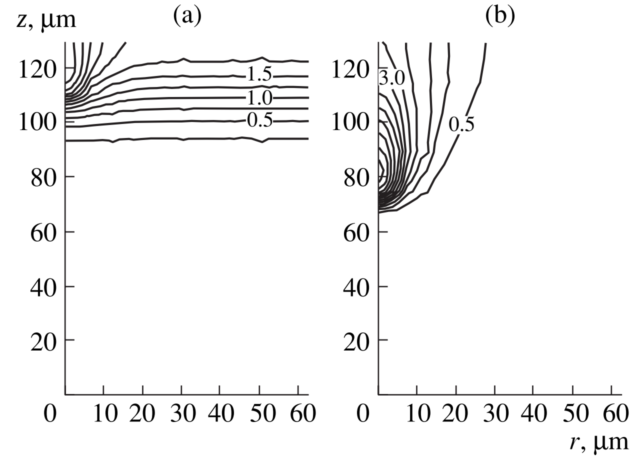


Fig. 1. Isotherms spaced at 500°C intervals (temperature values in 1000°C are shown by the curves): (a) I_0 = 1 MW/cm^2 and t = 80 μs (P = 0.5 MW, r_0 = 4 μm) and (b) I_0 = 4 MW/cm^2 and t = 210 μs (P = 4 MW and r_0 = 4 μm).

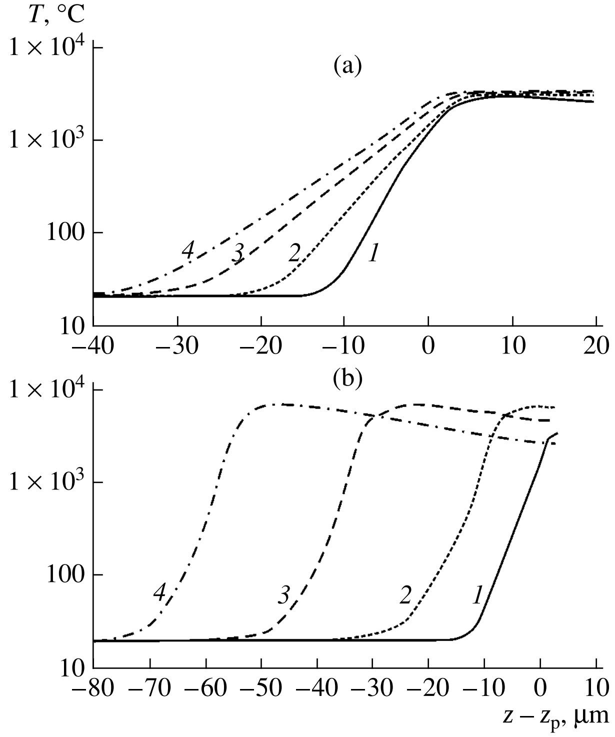


Fig. 2. Temperature distributions $T(t, z, r = 0)$ along the fiber at different time instants: (a) I_0 = 1 MW/cm^2, P = 0.5 MW, r_0 = 4 μm, and t = (*1*) 10, (*2*) 38, (*3*) 66, and (*4*) 80; (b) I_0 = 4 MW/cm^2, P = 4 MW, r_0 = 4 μm, and t = (*1*) 10, (*2*) 90, (*3*) 170, and (*4*) 210.

for $j = 1, \ldots, Nz$,

$$P^2_{1,j} = 0, \quad P^4_{1,j} = \frac{4K_{1,j}}{hr_1^2},$$

$$P^2_{Nr,j} = \frac{K_{Nr-\frac{1}{2},j}}{hr_{Nr-1}\left(r_{Nr} - r_{Nr-\frac{1}{2}}\right)}, \quad P^4_{Nr,j} = 0.$$

Thus, we arrive at the evolutionary problem

$$c_p(T)\rho(T)\frac{\partial}{\partial t}T(t, z, r)$$
$$= (\Lambda_z T)_{i,j} + (\Lambda_r T)_{i,j} + \alpha(T_{i,j})I_{i,j},$$

where

$$(\Lambda_z T)_{i,j} = P^1_{i,j}T_{i,j-1} + P^3_{i,j}T_{i,j+1} - Pz^5_{i,j}T_{i,j},$$

$$(\Lambda_r T)_{i,j} = P^2_{i,j}T_{i-1,j} + P^4_{i,j}T_{i+1,j} - Pr^5_{i,j}T_{i,j}.$$

Applying the method of splitting of spatial variables [18] to this problem, we construct an absolutely stable implicit scheme of first-order accuracy,

$$c_p(T^n_{i,j})\rho(T^n_{i,j})\frac{T^{n+\frac{1}{2}}_{i,j} - T^n_{i,j}}{\Delta t} = \left(\Lambda_r T^{n+\frac{1}{2}}\right)_{i,j},$$

$$c_p(T^n_{i,j})\rho(T^n_{i,j})\frac{T^{n+1}_{i,j} - T^{n+\frac{1}{2}}_{i,j}}{\Delta t} = (\Lambda_z T^{n+1})_{i,j} + \alpha(T^n_{i,j})I_{i,j},$$

where n and Δt are the number and length of the time step, respectively.

Each of the three-point difference equations obtained are easy to solve by the factorization (sweep) method.

RESULTS AND DISCUSSION

HAW propagation. Typical isotherms and temperature distributions along the fiber axis for the cases of low and high laser intensities are illustrated in Figs. 1 and 2, respectively. As expected, there is a temperature peak propagating toward the laser beam. After the peak, the temperature declines, since absorption in a layer α_p^{-1} = 18 μm thick and cooling via heat conduction decrease the laser intensity. At a lower intensity, cooling is stronger and the peak is more distinct.

The velocity v_f of the HAW front was determined from the time dependence of the front coordinate $z(t)$, which, in turn, was found from the equality $T(t, r = 0, z_f) = T_p$ (Figs. 2, 3). The dependence $z_f(t)$ was closely approximated by the linear function $z_f(t) = v_f t$ + const, and v_f was found from its slope. The calculation results, along with the experimental data [15], are presented in Fig. 4.

The HAW velocity. In combustion theory, the flame propagation rate is proportional to the square root of the specific power of energy release [19]. The corresponding expression, which is also used in discharge propa-

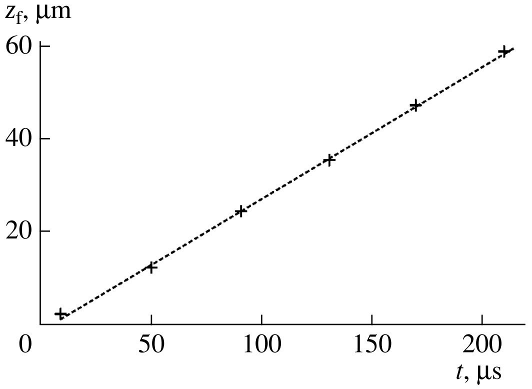


Fig. 3. Time dependences of the heat wave front coordinate z_f that is defined as the point where $T(t, z_f, r = 0) = T_p = 2000$ K. $I_0 = 4$ MW/cm^2, $P = 4$ MW, and $r_0 = 4$ µm. The dotted line depicts the dependence $z_f(t) = v_f t$ + const, where $v_f = 0.29$ m/s.

gation theory [2], in our case, has the form

$$v_f = \sqrt{k\alpha_p I_0/(T_p - T_0)}/(\rho c_p). \tag{3}$$

A slightly different expression for the HAW front velocity was proposed in [7, 8],

$$v_f = v_{f0}(\sqrt{(I_0/I_{ch}) + 1} - 1), \tag{4}$$

where $v_{f0} = k\alpha_p/2\rho c_p$ characterizes the linear growth of the front velocity with laser intensity in weak fields and $I_{ch} = k\alpha_p(T - T_0)/4$ is the intensity at which the linear dependence changes to the root one. For the calculation parameters in Fig. 4, we have $v_{f0} = 0.034$ m/s m/s and $I_{ch} = 5.8 \times 10^{-3}$ MW/cm^2.

At $I_0 \gg I_{ch}$, formula (4) coincides with (3). In contrast to the latter, formula (4) includes the energy spent on heating the mixture. The results of our calculation of front velocity v_f in the absence of cooling agree with the calculations by formula (4) when the intensity considerably exceeds the threshold (see below). However, expressions (3) and (4) yield close results (Fig. 4) under the conditions considered. The effect of cooling turns out to be more significant.

Threshold intensities. We determined threshold intensity I_{th} for various values of r_0 (Fig. 5). At $I_0 < I_{th}$, the heat wave is absent. An estimate of the threshold intensity, I_{th1}, can be found by equating the absorbed power $\alpha_p I_{th1}$ with the effective heat removal $(6(T_p - T_0)k)/r_0^2$. As a result, we arrive at

$$I_{th1} = \frac{6(T_p - T_0)k}{r_0^2 \alpha_p}. \tag{5}$$

In our calculation, the threshold was estimated for a velocity of 0.1 m/s. The values given by formula (5), our estimates, and the threshold intensities measured in [12] are compared in Fig. 5. It is seen that, for the parameter values considered, estimator (5) yields values that are roughly twice those obtained by the numerical calculation. At the same time, the calculation is in good agreement with the experimental data.

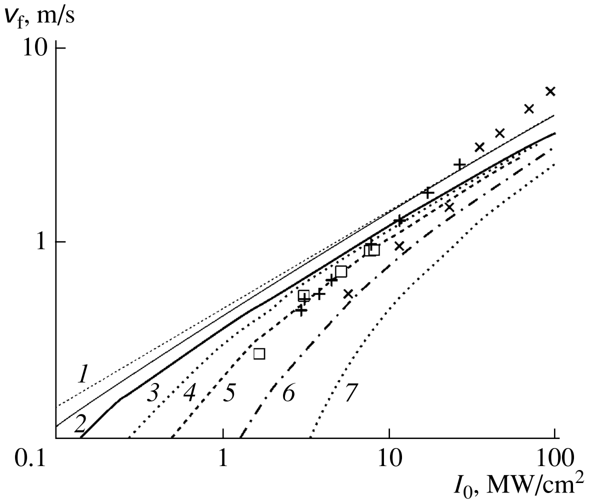


Fig. 4. HAW front propagation velocity vs. the laser intensity. Curves (*1*) and (*2*) show the results of calculation by formulas (3) and (4), respectively; curves (*3*)–(*7*) are the numerical results for core radius r_0 = 50, 20, 10, 4, and 2 µm, respectively. The symbols are data points for SiO$_2$–GeO$_2$ fibers with an outer radius of 125 µm [12]. The difference Δn between the refractive indices at the axis and periphery of the fiber is (×) 0.04, (+) 0.009, and (□) 0.0015; core diameter d = (×) 3.3, (+) 5.75, and (□) 11.05 µm.

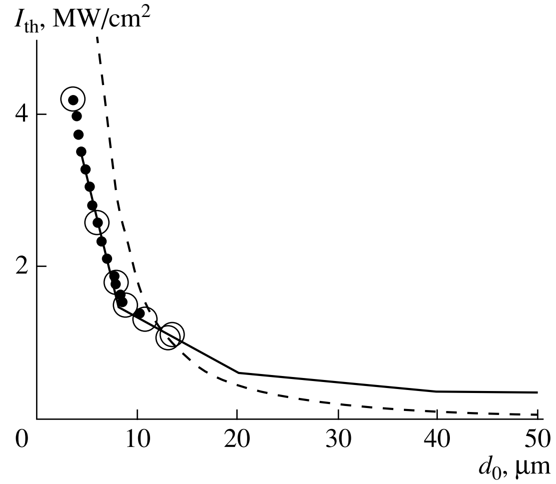


Fig. 5. Threshold laser intensity vs. core diameter d_0. The symbols are data points [12]; the solid line, numerical calculation; and the dashed line, estimate by formula (5).

CONCLUSIONS

Our calculations demonstrate that a heat absorption wave traveling along a fiber core can be fairly accurately described by a nonstationary 2D equation of heat conduction combined with a stationary equation for laser intensity. For intensities far exceeding a threshold, the HAW velocity is closely approximated by a well-known formula from combustion theory, i.e., is proportional to the square root of the radiation intensity. The threshold value can be estimated (in order of magnitude) by comparing the amount of heating with that of heat removal. The analytical value of the threshold is in good agreement with the experimental value.

The analytical wave velocity versus laser intensity dependences and the threshold intensities are consistent with the available experimental data.

ACKNOWLEDGMENTS

We thank E.M. Dianov for the statement of the problem, A.S. Biryukov and I.A. Bufetov for fruitful discussions, and A.A. Frolov for the digitization of the results.

REFERENCES

1. G. A. Askar'yan, A. M. Prokhorov, G. F. Chanturiya, and G. P. Shipulo, Zh. Éksp. Teor. Fiz. **44**, 2180 (1963) [Sov. Phys. JETP **17**, 1463 (1963)]; A. M. Prokhorov, *Quantum Electronics: Collection of Selected Works* (Atomizdat, Moscow, 1996), pp. 87–90 [in Russian].
2. N. V. Zelikin, N. E. Kask, V. V. Radchenko, *et al.*, Pis'ma Zh. Tekh. Fiz. **4**, 1296 (1978) [Sov. Tech. Phys. Lett. **4**, 522 (1978)].
3. N. E. Kask and L. S. Kornienko, Izv. Akad. Nauk SSSR, Ser. Fiz. Mat. Nauk **46** (6) (1982).
4. Yu. P. Raizer, *Laser-Induced Discharge Phenomena* (Nauka, Moscow, 1974; Consultants Bureau, New York, 1977).
5. R. Kashyap, in *Proceedings of the International Conference "Lasers-87," Lake Tahoe, 1987*, pp. 859–866.
6. R. Kashyap and K. J. Blow, Electron. Lett. **24**, 47 (1988).
7. D. P. Hand and P. St. J. Russel, Opt. Lett. **13**, 767 (1988).
8. D. P. Hand and P. St. J. Russel, in *Proceedings of the 14th European Conference on Optical Communication (ECOS), Brighton, 1988*, Part 1, pp. 111–114.
9. R. Kashyap, A. Sayles, and G. F. Cornwell, Proc. SPIE **2965**, 586 (1997).
10. T. J. Driscoll, J. M. Calo, and N. M. Lawandy, Opt. Lett. **16**, 1046 (1991).
11. E. M. Dianov, V. M. Mashinsky, V. A. Myzina, *et al.*, Sov. Lightwave Commun., No. 2, 293 (1992).
12. E. M. Dianov, I. A. Bufetov, A. A. Frolov, *et al.*, Kvantovaya Élektron. (Moscow) **32**, 476 (2002).
13. D. D. Davis, S. C. Mettler, and D. J. DiGiovanni, Proc. SPIE **2714**, 202 (1997).
14. D. D. Davis, S. C. Mettler, and D. J. DiGiovanni, Proc. SPIE **2966**, 592 (1997); Opt. Lett. **28**, 974 (2003).
15. A. P. Babichev, N. A. Babushkina, A. M. Bratkovskii, *et al.*, in *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meilikhov (Énergoatomizdat, Moscow, 1991; CRC, Boca Raton, 1997).
16. N. N. Kalitkin, *Numerical Methods* (Nauka, Moscow, 1978) [in Russian].
17. Yu. I. Sytsko and S. I. Yakovlenko, Laser Phys. **6**, 989 (1996).
18. G. I. Marchuk, *Methods of Splitting* (Nauka, Moscow, 1988) [in Russian].
19. Ya. B. Zel'dovich, *Chemical Physics and Hydrodynamics: Selected Works* (Nauka, Moscow, 1984), pp. 226–232; Ya. B. Zel'dovich and D. A. Frank-Kamenetskii, Zh. Fiz. Khim. **12**, 100 (1938).

Translated by A. Sidorova

Technical Physics, Vol. 50, No. 2, 2005, pp. 237–248. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 99–109.
Original Russian Text Copyright © 2005 by Bad'in, Zyuzin, Safronov, Slavnov, Tarasov.

ELECTRON AND ION BEAMS, ACCELERATORS

Computer Analysis of the Source Image Structure in 3D Electron–Optic Systems: I. An Electrostatic System

L. V. Bad'in, L. N. Zyuzin, S. I. Safronov, Yu. K. Slavnov, and R. P. Tarasov
Research Institute of Pulsed Technology, Moscow, 115304 Russia
e-mail: crabro@aport.ru

Received February 24, 2004; in final form, June 25, 2004

Abstract—Computer analysis of the arbitrary source image obtained in 3D electron–optic systems is performed. The systems involve electrostatic fields focusing and deflecting electron beams. Specifically, the structure of a net electron beam from an extended source at the crossover is examined. It is shown that the spread function of the source, which characterizes the imaging quality of the system most fully, may serve as a primary computational criterion. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

Fast processes, such as plasma and ballistic processes, thermonuclear fusion, etc., are basically studied with pulsed electron–optic converters, among which instruments offering a subpicosecond time resolution and introducing minor image distortions are attracting increasingly more attention. In analytical models of electron–optic systems (EOSs), the imaging quality is usually characterized by the shape of a spot of confusion in an appropriate cross section of the electron beam coming from a point source (emitter) and by the current density distribution over this spot (the spread function of a point source) [1]. However, the analytical studies are basically restricted to paraxial beams in axisymmetric EOSs. This means that the point-spread function cannot be considered as a measure of resolution of the instrument in terms of an analytical model. The fact is that, on the one hand, wide (instead of paraxial) beams are currently used according to modern technology's requirements and, on the other hand, deflecting plates and diaphragms used in the system break its axial symmetry. Under such conditions, only direct computer simulation of beam-forming and beam-deflecting fields followed by wide beam trajectory analysis can adequately describe EOS operation.

Today, preference is given to two types of EOSs with a subpicosecond resolution. In the former case, both focusing and scanning of the image are accomplished by means of an electric field alone. In the latter, a magnetic lens focuses the image, while an electric field scans. A great number of electrodes and their various configurations necessitate the solution of 3D field problems, which are difficult to solve even in the steady-state case. Moreover, the diameter of the spot of confusion in the focusing systems may be several orders of magnitude smaller than the characteristic type of the instrument. In the EOS design that is considered below, the cathode-to-screen distance (the longitudinal dimension of the converter) is 300 mm and the transverse dimension of the converter is 60 mm. At the same time, an elementary electron beam may be constricted to a circle of diameter $d = 0.01–0.20$ mm in the screen plane (the diameter of the circle depends on the point emitter position and the scanning potential). Therefore, finding the spread function, which is defined just on the spot of confusion, requires an exact calculation of many beam electron trajectories in the 3D electromagnetic field with the accuracy of solution of the field problem matched to the accuracy of trajectory analysis. Note also that the domain of definition of the point-spread function in the computer model is discrete and so must be smoothed by digital techniques of processing 2D discrete signals. Therefore, engineers employing computer models usually discard studying the spread function in favor of the resolution of the instrument, which is a much rougher criterion incompletely meeting the present-day requirements and inadequately reflecting the structure of the image.

In part 1 of this work, we study the imaging quality in typical multielectrode EOSs operating in the electrostatic regime. The second part of this work is devoted to 3D EOSs, where image transfer is accomplished by means of a stationary electromagnetic field.

BASIC DIAGRAM OF AN ELECTRON–OPTIC SYSTEM

The EOS efficiency depends on a variety of physical processes, each being responsible for the image structure in one way or another. It therefore seems reasonable to consider the most significant processes specifying the image structure in most EOS designs. Such processes are certainly image scanning, diaphragming, and adjustment of the instrument.

Figure 1 projects the basic diagram of a particular EOS design onto two orthogonal planes containing the

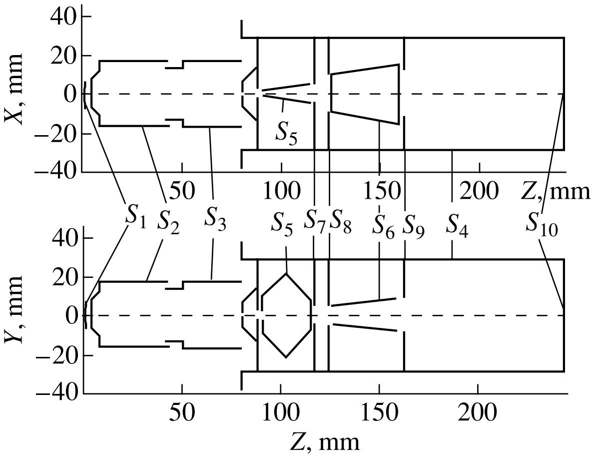


Fig. 1.

optical axis. Here, S_1 is the cathode; S_2, accelerating electrode; S_3, focusing electrode; S_4, anode chamber; S_5, deflectors; S_6, shutters; S_7–S_9, slit diaphragm; and S_{10}, screen. We assume that all electrodes S_i ($i = 1$–10) are conducting and are under given potentials. In this case, an electrostatic field is induced in the EOS.

COMPUTER MODEL OF AN ELECTRON–OPTIC SYSTEM WITH AN ELECTROSTATIC FIELD

As applied to EOSs of a general type [2–5], fast and exact numerical solution algorithms may be constructed in terms of the finite-group method (FGM), which exploits the local symmetries of the system's subsystems and iterative joining. The basic versions of the FGM are presented in [6–16]. In addition, for electrostatic-field (hereafter, electrostatic) EOSs, efficient algorithms of trajectory analysis that are based on adaptive (beam-attending) spatial computational grids [17] have been developed [2, 3]. Therefore, we will highlight only basic points in describing a computer model of electrostatic EOSs.

Consider set $\{S_i\}$ of N conducting screens S_i(S_i are the electrodes of a desired EOS) under the assumption that $\{S_i\}$ constitute multiply connected surface S in three-dimensional space R^3,

$$S = \bigcup_{i=1}^{N} S_i, \quad S_i \cap S_j = \varnothing, \quad i \neq j, \tag{1}$$

and that each screen j ($j = 1, 2, \ldots, N$) is under a given potential. Then, if $E = \{E_i\}$ ($i = 1, 2, 3$) is the vector of the electric field induced by the set S of screens, we have $E_i(x) = -\partial\varphi(x)/\partial x_i$, where $x = \{x_i\}$ are the Cartesian coordinates of point x in R^3 and $\varphi(x)$ is the potential of field E that satisfies the Laplace equation subject to the Dirichlet conditions on S:

$$\begin{aligned} &\Delta\varphi(x) = 0, \quad \Delta \equiv \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2 + \partial^2/\partial x_3^2, \\ &x \in R^3\backslash S, \\ &\varphi(x) = \sum_{i=1}^{N} \chi_{S_i}(x) V_i, \quad x \in S. \end{aligned} \tag{2}$$

Here, $\chi_{S_i}(x)$ is the characteristic function of a set of points on surface S_i and Δ is the Laplacian.

A solution to boundary-value problem (2) may be represented as the potential

$$\varphi(x) = \int_S \frac{u(x_1)}{|x - x_1|} d\mu(x_1), \quad x_1 \in R^3\backslash S$$

of a simple fiber with density $u(x)$, $x \in S$, that satisfies the boundary integral equation of the first kind

$$\begin{aligned} &[Au](x) = f(x), \quad [Au](x) = \int_S \frac{u(x_1)}{|x - x_1|} d\mu(x_1), \\ &f(x) = \sum_{i=1}^{N} \chi_{S_i}(x) V_i, \end{aligned} \tag{3}$$

where $d\mu(x)$ is the contraction of an element of the Euclidean space onto surface S and $|x - x_1|$ is the Euclidean distance between points x and x_1.

A numerical solution to Eq. (3) is constructed in terms of the FGM version that employs finite groups of symmetries discontinuously acting on surfaces S_i and takes into account the invariance of operators $A_{ii} = \chi_{S_i}(x) A \chi_{S_i}(x)$ (the contraction of A on S_i) under transformation of these groups.

The characteristics of EOSs, including the point-spread function, are determined by trajectory analysis. Namely, the electron trajectories in the electrostatic field are described through the Cauchy problem for a set of ordinary differential equations

$$\frac{d}{dt}\left(m\frac{dr}{dt}(x, t)\right) = -e\nabla\varphi(x), \tag{4}$$

where r is the radius vector of a particle with mass m and charge e and t is the transit time.

In turn, the initial conditions of particle escape distinguish the initial velocities of the electrons emitted from a point by magnitude and direction. Such a distribution, characterizing a specific cathode, is usually found experimentally.

Let P be a point from which electrons leave a cathode (point emitter). It is routinely assumed in simulation that the probability of a particle escaping within a solid angle $d\omega$ at an angle θ to normal n_p varies as $\cos^k\theta d\omega$, $k = 1, 2, 3, \ldots$ ($k = 1$ corresponds to the Lambert law). Unless otherwise stated, we below consider bundles of N equiprobable trajectories obeying the angular distribution law $\cos^3\theta$ on assumption that the escape energy of the electrons is the same (the energy spread in the source is disregarded).

The properties of the image will be studied on a particular surface s, where s may be a screen, image surface, or given secant plane, with the electrons emitted from point P on the cathode being gathered within a segment ds_p on surface s.

Let (τ^1, τ^2, τ^3) be a set of orthogonal coordinates that is chosen in such a way that surface s is a part of coordinate surface $\tau^3 = \text{const}$. In this case, each of the electron trajectories $\{\tau_i\}$, $i = 1, 2, \ldots, N$, emerging from cathode point P on surface s is described by coordinates $\{(\tau_i^1, \tau_i^2)\}$. Then, the spreading center $\bar\tau = (\bar\tau^1, \bar\tau^2)$ of an elementary beam and the rms deviation σ (or $\sigma^{(\tau^k)}$, $k = 1, 2$, in coordinate-wise form) from the center of gravity are given by

$$\bar\tau^k = \frac{1}{N}\sum_{i=1}^{N}\tau_i^k, \quad \sigma^{(\tau^k)} = \left(\frac{1}{N}\sum_{i=1}^{N}(\tau_i^k - \bar\tau^k)^2\right)^{1/2}; \tag{5}$$
$$k = 1, 2; \quad \sigma = \left(\frac{1}{N}\sum_{i=1}^{N}|\tau_i - \bar\tau|^2\right)^{1/2},$$

where $|\tau_i - \bar\tau|$ is the Euclidean distance between points τ_i and $\bar\tau$ ($\tau_i, \bar\tau \in s$).

The rms deviation from the center of gravity gives an estimate (rather accurate in the case of axisymmetric EOSs) of only the size of the spot of confusion (segment ds_p). The shape and structure of this spot for an elementary electron beam can be judged from the spread function $S_p(\tau)$ for point emitter P, which is determined as follows:

$$S_p(\tau) = \begin{cases} n, & \tau \in \{\tau_{k_1} = \ldots = \tau_{k_n}\} \\ 0, & \tau \notin \{\tau_i\}, \quad i = 1, 2, \ldots, N, \end{cases} \quad \tau \in s. \tag{6}$$

Point-spread function $S_p(\tau)$ is a function of continuous argument $\tau \in s$ with the discrete domain of definition. Such functions are inconvenient for analysis and visualization. Instead of $S_p(\tau)$, we will consider approximating functions $S_p(i, j)$. To construct them, we insert the spot of confusion ds_p ($S_p(\tau) \equiv 0$ at $\tau \notin ds_p$) into a rectangle ds_p the center of which coincides with the spread center $\bar\tau$ of electron beam $\{\tau_i\}$. This rectangle is covered by a uniform rectangular grid

$$\Omega(\{M_1\}, \{M_2\}) = \bigcup_{i, j \in \{\{M_1\}, \{M_2\}\}} \omega_{ij}, \tag{7}$$

where $\{M_1\}$ and $\{M_2\}$ are sets of integers that are taken so that the spread center $\bar\tau$ of the beam is the center of the rectangular mesh ω_{00}.

Assigning the values equal to the number of electrons falling into a mesh ω_{ij} to $S_p(i, j)$, we obtain the desired approximating function.

Spread functions $S_p(i, j)$ characterize the spatial structure of an elementary electron beam fairly adequately. However, the entire information on the beam structure that is contained in these functions may be extracted by applying smoothing digital filtering. The essence of this procedure will be understood by visualizing the values of some of the spread functions.

Below, the digital processing of spread functions $S_p(i, j)$ uses separation adaptive digital filters of 2D signals that are based on discrete B splines. Smoothing filtering is carried out in the interactive mode: first, a discrete scaling B spline is prescribed and then sharpening digital iterative filtering is accomplished. Earlier, such a procedure of smoothing 2D pulsed signals was applied in computer analysis of inverse multidimensional problems of unsteady heat conduction [18]. Hereafter, the digitally processed spread functions $S_p(i, j)$ are designated as $\hat S_p(i, j)$.

When numerically analyzing the structure of images transferred by EOSs, we considered not only elementary but also net beams, which comprise all the electrons emitted from a given set $\{P_k\}$ of cathode points. By analogy with elementary beams, net beams are also characterized by appropriate spread functions $S(\tau)$, $\hat S(i, j)$, and $S_p(i, j)$.

It should be noted in conclusion that spread functions $S_p(\tau)$ ($S_p(i, j)$) characterize the spatial structure of an electron beam and can be used to estimate only the spatial (technical) resolution of the instrument. However, the electron trajectories in our EOS model are functions of the coordinates and time. Consequently, we may consider functions $S_p(\tau, t)$ ($S_p(i, j, t)$) instead of $S_p(\tau)$ ($S_p(i, j)$) and thus obtain an estimate of the time resolution of the instrument. This issue calls for special investigation.

IMAGE STRUCTURE IN AXISYMMETRIC EOSs

Computer analysis of the image structure in the 3D EOSs under consideration is a preliminary to numerical simulation of an axisymmetric EOS consisting of electrodes S_1–S_4 (Fig. 1). Electrodes S_5–S_9 inside anode chamber S_4 are omitted from consideration.

A computer model of axisymmetric EOSs that is based on numerical solution algorithms for one-dimensional boundary integral equations from the potential theory (these algorithms provide a high accuracy of constructing axisymmetric electrostatic fields) and on trajectory analysis methods could be elaborated and program-implemented only after the work by Antonenko [19] had been published (1964). The early routines of this type appeared in the United States in the mid-1970s.

Below, we give the results of computer analysis of the real axisymmetric EOS design (see above) that are necessary for further discussion.

Numerical simulation data for axisymmetric EOSs as representatives of multielement three-dimensional structures (Fig. 1) may also help in optimizing computer analysis of these structures. Numerical experiments have shown that the external electric field generated by electrodes S_1–S_4 penetrates into anode chamber S_4 insignificantly. In turn, the electric field inside the anode chamber, which is generated by electrodes S_5 and S_6, is damped near the circular hole of chamber S_4. In other words, the interaction between these fields is weak and local. This makes it possible to perform a comparative calculation using two program packages intended for simulating axisymmetric and 3D EOSs. Such a calculation greatly simplifies the selection of a set of boundary elements on desired boundary surface S formed by all the electrodes of a 3D EOS and provides effective control of the 3D simulation accuracy. To do this, the structure formed by electrodes S_1–S_4 is simulated using the program package for analysis of axisymmetric systems. When trajectory analysis is carried out, the number of calculation points on the generatrix of the surface of revolution S_v, $S_v = \bigcup_{i=1}^{4} S_i$, and the mesh spatial size in the attendant grid are selected so as to fit a given calculation accuracy ε (e.g., ε = 0.1%). Then, using the program package for analysis of 3D EOSs, the desired EOS design (Fig. 1) is simulated with electrodes S_5–S_9 grounded (i.e., these electrodes are under the potential of anode chamber S_4). Here, a set of boundary elements on surface S, $S = \bigcup_{i=1}^{9} S_i$, and an appropriate attendant grid are selected in such a way that the results of simulation using the two packages coincide within an error not exceeding ε. If necessary, the set of boundary elements and the attendant grid are refined by applying low scanning potentials on electrodes S_5 and S_6. It should be emphasized that results of simulation must include all computable EOS parameters and characteristics (from the magnification to the point-spread function).

(i) Crossover. In focusing EOSs, the net electron beam (i.e., the beam of the electrons emitted from all cathode points) has a minimal cross-sectional area (crossover) in some plane (the plane of crossover). In the analytical models of EOSs, much attention is paid to the crossover position and the current density in it. For axisymmetric EOSs, it was found [1] that the current distribution over the crossover is Gaussian. Analytical and experimental data of this sort are of great value for EOS designers.

Consider cylindrical coordinate system (r, ψ, z) that is related to Cartesian system XYZ depicted in Fig. 1. Let the net electron beam consist of six elementary

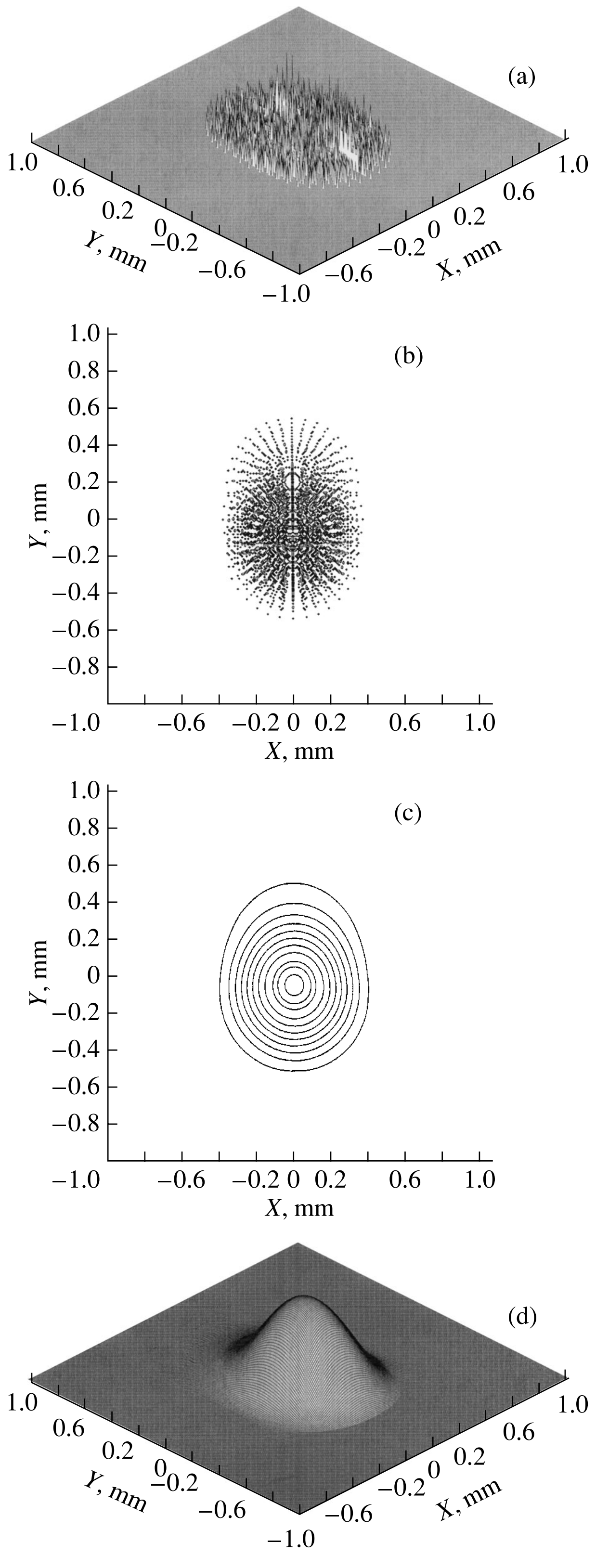


Fig. 2.

beams issuing out of cathode points (r_k, $\psi = \pi/2$, $z = 0$), where $r_k = kmm$ and $k = 0, 1, \ldots, 5$; that is, all emitters (emitting points) lie on the positive Y semiaxis. It is also assumed that each of the elementary beams (the electron escape energy $U_0 = 0.35$ eV) may trace 512 equiprobable trajectories with the escape angle distribution law $\cos^3\theta$. The potential difference between the cathode (S_1) and anode chamber (S_4) is 15 kV.

The values of the spread function $S(i, j)$ of the net beam in the crossover are visualized in Fig. 2a and its isolines, in Fig. 2b. The visualization of the spread function (Fig. 2a) seems to carry little information, while Fig. 2b, which plots the $S(i, j)$ isolines, gives insight into the shape and sizes of the spot of confusion of the net beam in the crossover (the maximal size of the spot is 0.8×1.0 mm). In addition, it is clearly seen from Fig. 2b that the elementary beams are not completely coincident in the Y direction (complete coincidence takes place in EOSs with magnetic focusing, where the crossover area is larger; see part II of this work). However, neither the plot of function $S(i, j)$ (Fig. 2a) nor its isolines (Fig. 2b) give a comprehensive idea of the net beam structure in the crossover.

The structure of the net beam in the crossover became evident after function $S(i, j)$ had been digitally processed by means of adaptive B-spline-filters. Figures 2c and 2d show, respectively, the isolines and values of function $\hat{S}(i, j)$ (function $S(i, j)$ after digital processing).

Since the cathode points generating the net beam under consideration lie on the same ray passing through the center of the cathode and a solution to the problem stated is invariant under transformations from rotation group C_∞, the plot of $\hat{S}(i, j)$ (Fig. 2d) lets us conclude that the digitally processed point-spread functions of the net beams emitted from cathode points (r_n, ψ_m, $z = 0$) (where $r_n = nR/N$; $\psi_m = 2\pi(m-1)/M$; $n = 1, 2, \ldots, N$; $m = 1, 2, \ldots, M$; and R is the radius of the cathode) will have the bell-shaped (near-Gaussian) form in the crossover. This assertion was confirmed experimentally.

(ii) Structure of elementary beams. The behavior of the net beam in the working region of an axisymmetric EOS sheds light not only on the crossover parameters but also on other EOS characteristics of applied interest. Among them is, in particular, the image surface, since the centers of gravity $\bar{\tau}_i$ ($i = 0, 1, \ldots, 5$) of the focused elementary beams constituting the given net beam lie on the generatrix of the image surface. The potentials of electrodes S_2 and S_3 (the potentials of S_1 and S_4 are fixed, see above) are taken such that the elementary beam that escapes from the cathode with the radial coordinate $r_2 = 2$ mm is focused on the screen. Note that the image surface is approximated fairly well by an ellipsoid of revolution and that the beams with radial escape coordinates $r < (>)\ r_2$ are focused behind (before) the screen plane. This fact is embodied in the respective spread functions.

Figure 3a shows the values (on the left) and isolines (on the right) of the spread function $S_{p_0}(i, j)$ of point emitter $P_0 = P$ ($r_0 = 0$, $\psi = \pi/2$, $z = 0$) on the screen. Figure 3b shows the same for the spread function $S_{p_2}(i, j)$ of point emitter $P_2 = P$ ($r_2 = 2$ mm, $\psi = \pi/2$, $z = 0$); Fig. 3c, for the spread function $S_{p_4}(i, j)$ of point emitter $P_4 = P$ ($r_4 = 4$ mm, $\psi = \pi/2$, $z = 0$). Similarly, Figs. 4a–4c visualize the values and isolines of spread functions $\hat{S}_{p_0}(i, j)$, $\hat{S}_{p_2}(i, j)$, and $\hat{S}_{p_4}(i, j)$ (i.e., functions $S_{p_k}(i, j)$, where $k = 0, 2$, and 4, after adaptive B-spline digital filtering), respectively.

Only the central beam has a regular structure (Fig. 3), which is due to the initial conditions of electron escape from the cathode. Even a small offset of an emitting point (emitter) from the cathode center (by several hundredths of a millimeter) causes a rapid breakdown of the structure. At the same time, spread functions $\hat{S}_{p_k}(i, j)$ ($k = 0, 2$, or 4) on the image surface had a clearly cut bell-shaped form, which deforms as the emitter is displaced from the cathode center (Fig. 4b). The spread functions of the elementary beams that are focused behind the screen plane ($r < 2$ mm) appear on the screen as a flattened-top bell (Fig. 4a), while the bell-shaped spread functions of the beams focused before the screen plane ($r > 2$ mm) have a diffuse base (substrate) (Fig. 4c). Thus, in electrostatic axisymmetric EOSs, the structures of the elementary beams that are focused before and behind the image surface radically differ and the change takes place fairly rapidly. It should be noted that, in the magnetic focusing instruments where focusing is provided by a dc coil, the beams behave in the same manner (see part II of this work).

EFFECT OF SCANNING ON THE IMAGE STRUCTURE

Let us see how the fields scanning the image that is formed in an electrostatic EOS influence the structure of the elementary beams constituting the net beam considered in the previous section (these fields are generated by the potential applied to deflecting plates S_5, Fig. 1). Scan potential φ is applied to deflecting plates S_5 in two, one-phase (asymmetric) and two-phase (symmetric), modes. In the former case, one of the plates is under the (zero) potential of the anode chamber (i.e., is grounded). In the latter, the potentials applied to the plates are equal in magnitude but differ in sign (relative to the potential of the anode chamber). We will assume that the elementary beams being scanned are emitted from cathode points P_k ($x = 0$, $y_k = kmm$, $z = 0$; $k = 0, 1, \ldots, 5$) and consider the image structure on both the image surface and screen. The beam focus is

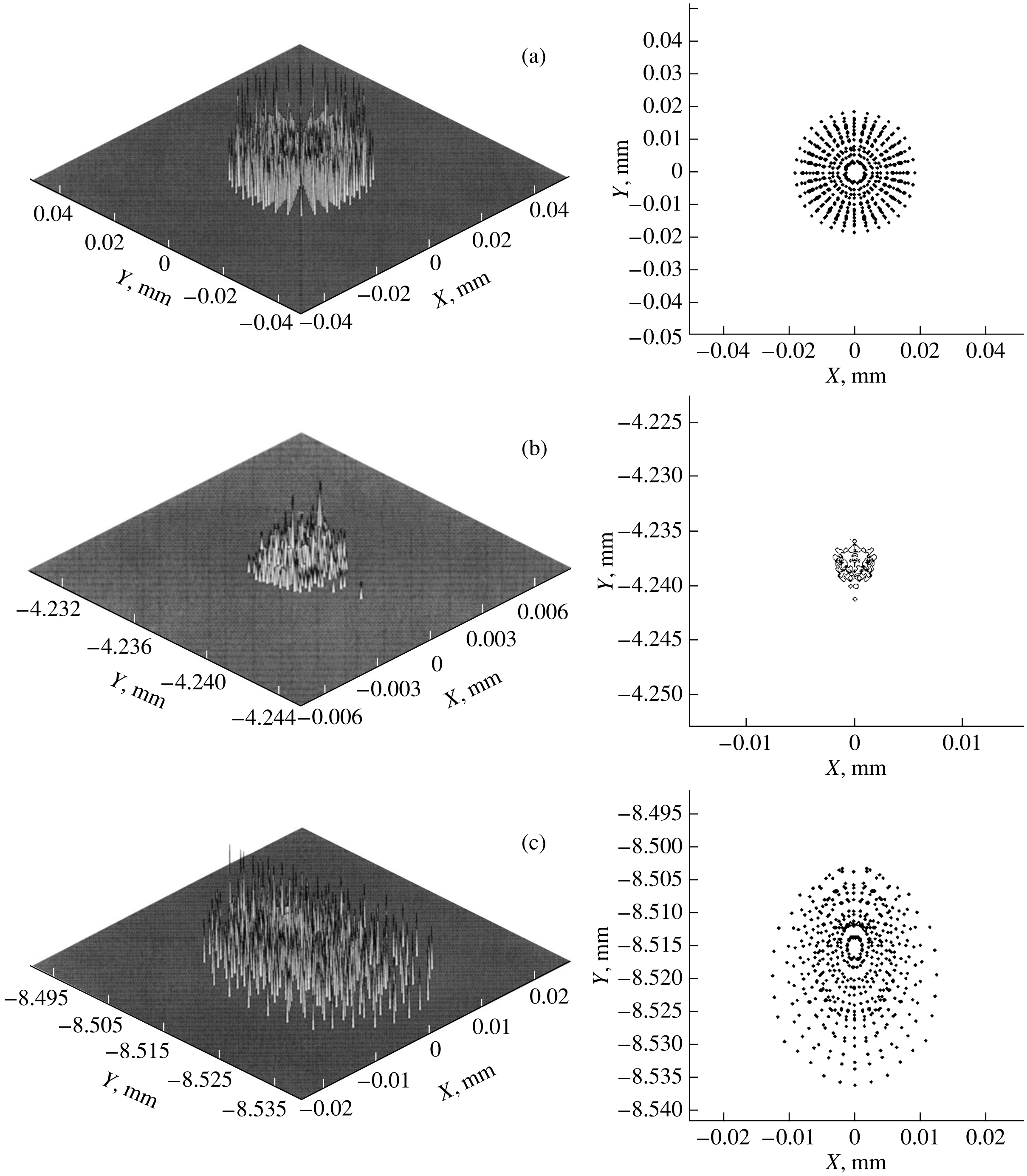


Fig. 3.

determined from the minimum of rms deviation $\sigma_\varphi(k)$ from the center of gravity $\bar{\tau}_\varphi(k)$. Then, we will analyze the dependence of spread functions $S^{(\varphi)}_{p_k}(i, j)$ $(\hat{S}^{(\varphi)}_{p_k}(i, j))$ on scan potential φ.

(i) For the symmetric scan mode and potentials $\varphi = \pm 0, \pm 100, \pm 200$, and ± 300 V, the values of rms deviation $\sigma_\varphi(k)$ and the Cartesian coordinates $(x^{(\varphi)}_k, y^{(\varphi)}_k, z^{(\varphi)}_k)$ of the centers of gravity $\bar{\tau}_\varphi(k)$ $(k = 0, 1, \ldots, 5)$ of the elementary beams on the image surface are given (in millimeters) in Table 1.

Note that, for scan potentials $\varphi = \mp 0, \mp 100, \mp 200$, and ∓ 300 V, the quantities listed in Tables 1 and 2 remain virtually unchanged (for coordinate x, the scan direction, this is true up to sign).

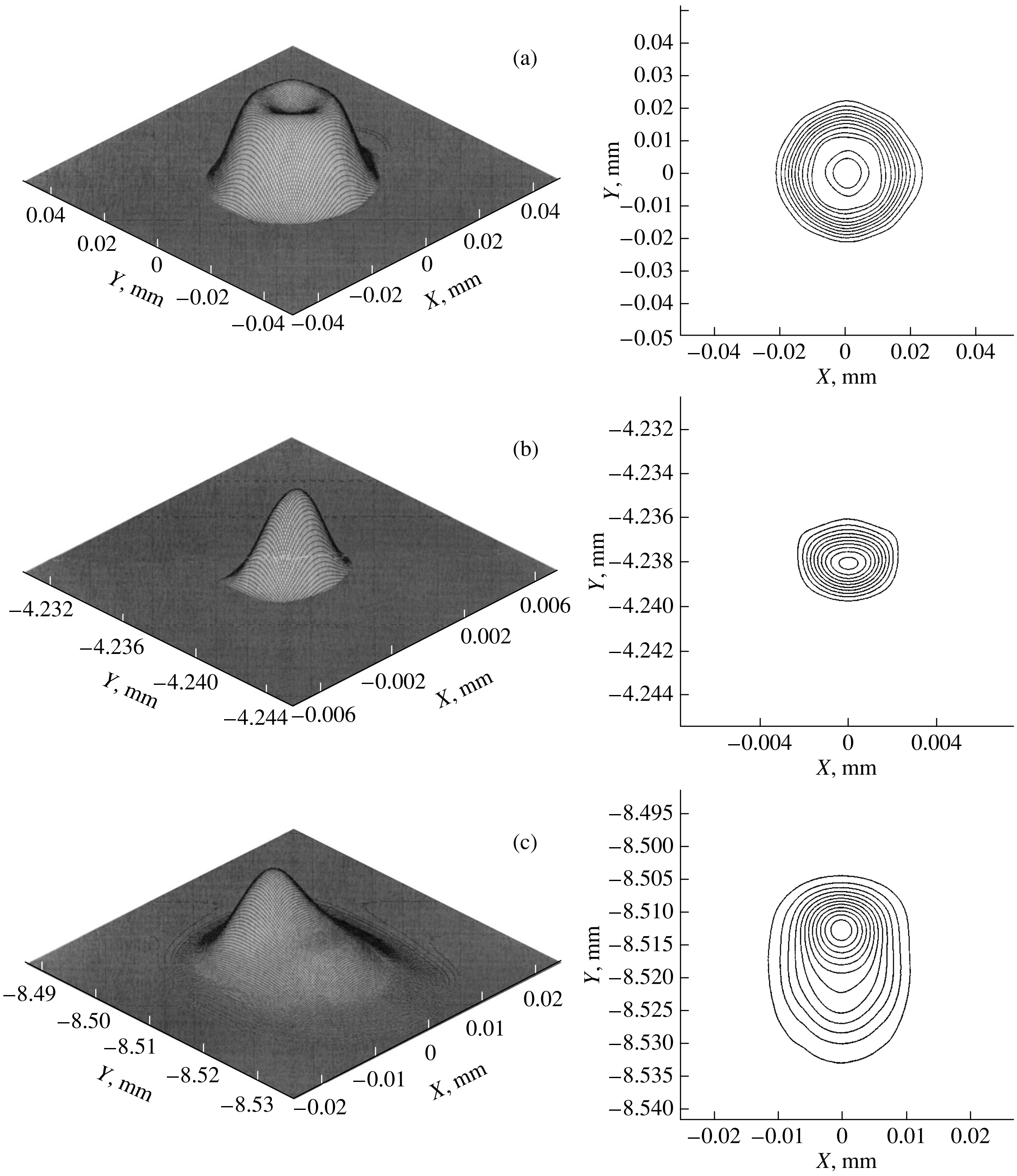


Fig. 4.

It follows from Tables 1 and 2 that the depth of the image surface along coordinate z (the optical axis of the system) is about 22.2 mm but the values of $\sigma_\varphi(k)$ for the elementary beams (i.e., for φ and k fixed) on the image surface and on the screen in general ($\varphi \neq 0$) vary insignificantly (a considerable depth of focus of the image). As the scan potential increases, the resolution of the instrument drops on both the image surface and screen (by a factor of 1.5–5 at the edge of the screen). Yet, the resolution of the structure simulated remains high in this scan mode.

Table 3 lists the values of rms deviation $\sigma_\varphi(k)$ and the Cartesian coordinates ($x_k^{(\varphi)}$, $y_k^{(\varphi)}$, $z_k^{(\varphi)}$) ($k = 0, 1, \ldots, 5$) of the centers of gravity of the elementary beams on the image surface for the one-phase scan mode with $\varphi =$

Table 1. Image surface (two-phase mode)

$\varphi \backslash k$		0	1	2	3	4	5
0 V	$\sigma_\varphi(k)$	0.004	0.019	0.013	0.001	0.003	0.003
0 V	$x_k^{(\varphi)}$	0	0	0	0	0	0
	$y_k^{(\varphi)}$	0	–2.22	–4.51	–6.63	–8.71	–10.90
	$z_k^{(\varphi)}$	255.5	254.9	257.1	253.2	250.5	250.4
+100 V	$\sigma_\varphi(k)$	0.007	0.019	0.015	0.007	0.008	0.007
–100 V	$x_k^{(\varphi)}$	5.98	5.97	6.1	5.95	5.82	5.82
	$y_k^{(\varphi)}$	0	–2.10	–4.30	–6.31	–8.27	–10.33
	$z_k^{(\varphi)}$	245.6	245.4	248.6	244.8	241.6	241.5
+200 V	$\sigma_\varphi(k)$	0.012	0.021	0.020	0.013	0.015	0.015
–200 V	$x_k^{(\varphi)}$	11.81	11.84	12.17	11.85	11.53	11.50
	$y_k^{(\varphi)}$	0	–2.09	–4.29	–6.30	–8.21	–10.23
	$z_k^{(\varphi)}$	243.7	244	248.1	244.2	240.1	239.6
+300 V	$\sigma_\varphi(k)$	0.024	0.03	0.032	0.027	0.026	0.025
–300 V	$x_k^{(\varphi)}$	16.68	16.69	17.37	16.85	16.40	16.33
	$y_k^{(\varphi)}$	0	–1.98	–4.13	–6.03	–7.86	–9.77
	$z_k^{(\varphi)}$	235.3	235.3	240.8	236.5	232.7	232

Table 2. Screen (two-phase mode)

$\varphi \backslash k$		0	1	2	3	4	5
±0 V	$\sigma_\varphi(k)$	0.013	0.022	0.020	0.009	0.006	0.006
	$y_k^{(\varphi)}$	0	–2.11	–4.24	–6.36	–8.49	–10.61
±100 V	$\sigma_\varphi(k)$	0.007	0.019	0.016	0.007	0.01	0.009
	$x_k^{(\varphi)}$	5.99	6	6	6	6	6.01
	$y_k^{(\varphi)}$	0	–2.11	–4.24	–6.35	–8.49	–10.61
±200 V	$\sigma_\varphi(k)$	0.012	0.021	0.02	0.014	0.017	0.017
	$x_k^{(\varphi)}$	11.99	12	12	12	12.01	12.02
	$y_k^{(\varphi)}$	0	–2.11	–4.24	–6.36	–8.50	–10.63
±300 V	$\sigma_\varphi(k)$	0.028	0.033	0.033	0.030	0.032	0.031
	$x_k^{(\varphi)}$	18	18.01	18.01	18.01	18.02	18.04
	$y_k^{(\varphi)}$	0	–2.12	–4.26	–6.38	–8.52	–10.66

Table 3. Image surface (one-phase mode)

$\varphi\backslash k$		0	1	2	3	4	5
200 V	$\sigma_\varphi(k)$	0.026	0.032	0.036	0.036	0.032	0.034
0 V	$x_k^{(\varphi)}$	6.38	6.21	5.72	5.72	5.91	5.88
	$y_k^{(\varphi)}$	0	–2.15	−4.01	−6.01	–8.25	–10.30
	$z_k^{(\varphi)}$	255.7	251.4	239.5	239.5	244	243
400 V	$\sigma_\varphi(k)$	0.066	0.069	0.078	0.079	0.074	0.076
0 V	$x_k^{(\varphi)}$	13.11	12.32	10.17	10.36	11.30	11.16
	$y_k^{(\varphi)}$	0	–2.12	–3.59	–5.46	–7.88	–9.72
	$z_k^{(\varphi)}$	260.5	250.7	224.4	226.6	237.9	236
600 V	$\sigma_\varphi(k)$	0.081	0.084	0.091	0.092	0.089	0.092
0 V	$x_k^{(\varphi)}$	17.86	15.97	12.95	13.2	14.73	14.46
	$y_k^{(\varphi)}$	0	–1.86	–3.12	–4.74	–6.96	–8.53
	$z_k^{(\varphi)}$	246.2	230.7	205.9	207.9	220.2	217.8

200, 400, and 600 V. The values of $\sigma_\varphi(k)$ on the screen are presented in Table 4.

For φ = –200, –400, and –600 V in the asymmetric mode, the associated values of $\sigma_\varphi(k)$ and ($x_k^{(\varphi)}$, $y_k^{(\varphi)}$, $z_k^{(\varphi)}$) differ from those listed in Tables 3 and 4 insignificantly.

The extent of the image surface along the z coordinate in this mode is much larger, ≈51.2 mm, than in the symmetric scan mode. As in the case of the two-phase mode, the values of $\sigma_\varphi(k)$ for a specific beam on the image surface and on the screen are comparable and the drop in the resolution with increasing scan potential is even higher: by a factor of 4–17 at the edge of the screen.

Tables 1 and 3 (as well as 2 and 4) imply that the two-phase scan mode is two to three times more efficient in terms of the imaging quality than the one-phase mode, at least in the EOS design under consideration. For the deflecting system in magnetic-focusing EOSs, the results are qualitatively the same (see part II of this work).

(ii) The rms deviation may serve to estimate only the area of the spot of confusion (for an elementary or net beam) and makes it possible to reveal the factors that have the most obvious effect on the image structure. The image structure itself is most fully characterized by the spread functions.

Figure 3 demonstrates the values and isolines of spread functions $S_{p_k}^{(\varphi)}(i, j)$ for point emitters $P_0 = P$ (x = 0, y = 0, z = 0), $P_2 = P$ (x = 0, y = 2 mm, z = 0), and $P_4 = P$ (x = 0, y = 4 mm, z = 0) when plates S_5 are grounded ($\varphi = \pm 0$). In Fig. 4, the values and isolines of functions $\hat{S}_{p_k}^{(0)}(i, j)$ (functions $S_{p_k}^{(0)}(i, j)$) after filtering) are shown. Consider the effect of the scan potential on the structure of these beams.

For the scan potential $\varphi = \pm 300$ V in the symmetric scan mode, Fig. 5 shows the isolines of spread functions $S_{p_k}^{(\pm 300)}(i, j)$ (k = 0, 2, 4) (left-hand column) and the values of $\hat{S}_{p_k}^{(\pm 300)}(i, j)$ (right-hand column). Figures 5a–5c correspond to emitters P_0, P_2, and P_4, respectively. Figures 6a–6c present the isolines of $S_{p_k}^{(600)}(i, j)$ (left-hand column) and the values of $\hat{S}_{p_k}^{(600)}(i, j)$ (right-hand

Table 4. Screen (one-phase mode)

$\varphi\backslash k$	0	1	2	3	4	5
$\begin{pmatrix}200\text{ V}\\0\text{ V}\end{pmatrix}$	0.029	0.032	0.038	0.038	0.032	0.034
$\begin{pmatrix}400\text{ V}\\0\text{ V}\end{pmatrix}$	0.069	0.069	0.085	0.083	0.075	0.077
$\begin{pmatrix}600\text{ V}\\0\text{ V}\end{pmatrix}$	0.081	0.086	0.116	0.11	0.096	0.1

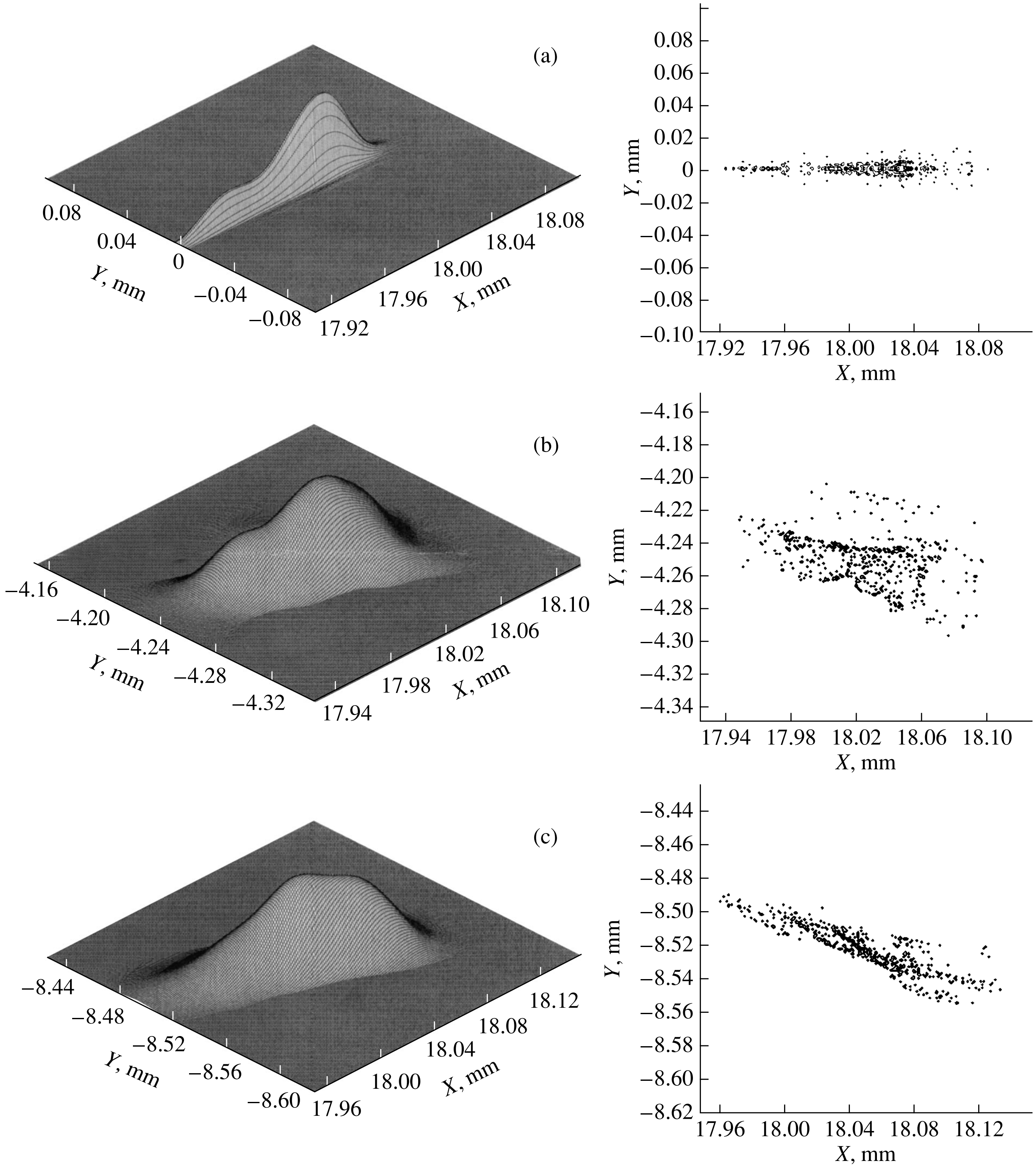


Fig. 5.

column) for point emitters P_k (k = 0, 2, 4) in the one-phase scan mode at φ = 600 V.

Figures 3–6 validate the conclusions regarding the rms deviation that the area of the spot of confusion in a given cross section of an elementary electron beam depends on the scan mode and potential, as well as on the point emitter position. However, the rms deviation characterizes only a circle (coordinate-wise an ellipse) containing most electrons of the beam but cannot inform us about the true shape and *a fortiori* structure of the spot of confusion. For example, the spot of confusion for the undeflected central beam is close to a circle of radius $r \approx \sigma_0(0) = 0.013$ mm (Figs. 3a, 4a). When the deflecting plates are under potential $\varphi = \pm 300$ V in the symmetric scan mode, the related spot of confusion on the screen is heavily extended along the scan direc-

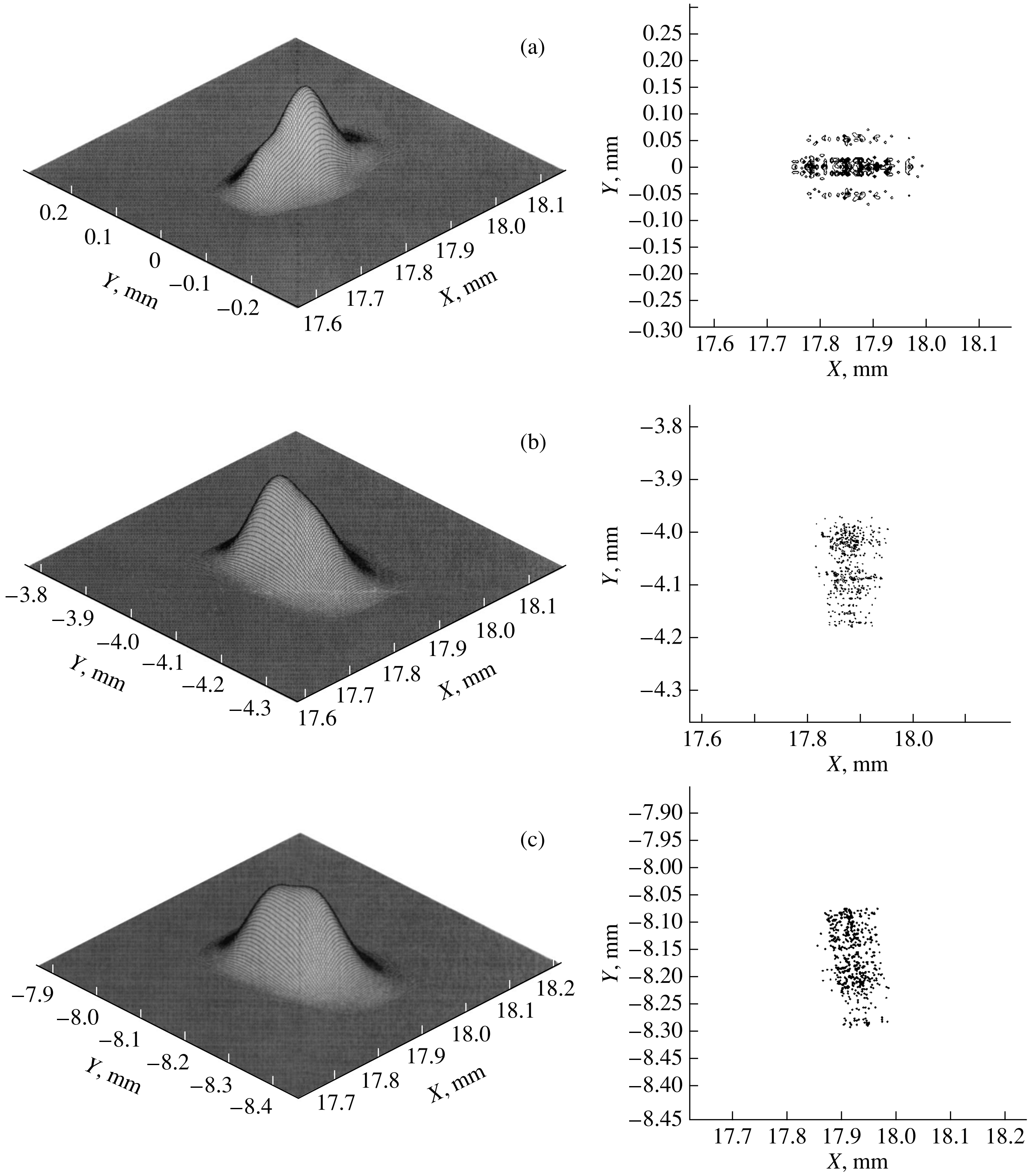


Fig. 6.

tion (Fig. 5a) (accordingly, $\sigma^{(x)}_{\pm 300}(0)$ = 0.028 mm and $\sigma^{(y)}_{\pm 300}(0)$ =0.003 mm). Such an ellipse with the semiaxis a = 0.028 mm and b = 0.003 mm, which is centered at the center of gravity of the beam, encircles about 70% of the trajectories. The same is also true for the other elementary beams being scanned (Figs. 3–6).

Detailed investigation of the effect of scan fields on the image structure in EOSs requires a large body of computing experiments. However, certain conclusions can be drawn by comparing the spread functions shown in Figs. 4–6. In the absence of the scan potential, spread functions $\hat{S}^{(0)}_{p_k}(i, j)$ (k = 0, 2, 4) for the elementary beams emitted from points P_0, P_2, and P_4 on the cathode (Fig. 4) substantially differ from each other on the screen, adequately reflecting the structure of the beams. Next, in the symmetric (Fig. 5) and asymmetric (Fig. 6) scan modes (the respective scan potentials are φ = ±300 and 600 V), the resolutions of the instrument at the edge of the screen are nearly the same in the scan direction, whereas the two-phase mode offers a much (several times) higher resolution in the perpendicular direction. Other features related to the beam structures (Figs. 4–6) are less obvious and call for further analysis.

REFERENCES

1. W. Glaser, *Grundlagen der Elektronenoptik* (Springer, Vienna, 1952; Gostekhizdat, Moscow, 1957).
2. S. K. Demin, S. I. Safronov, and R. P. Tarasov, Zh. Tekh. Fiz. **68** (2), 97 (1998) [Tech. Phys. **43**, 222 (1998)].
3. S. K. Demin, S. I. Safronov, and R. P. Tarasov, Zh. Tekh. Fiz. **68** (7), 126 (1998) [Tech. Phys. **43**, 861 (1998)].
4. S. I. Safronov and R. P. Tarasov, Zh. Tekh. Fiz. **69** (6), 1 (1999) [Tech. Phys. **44**, 609 (1999)].
5. S. I. Safronov and R. P. Tarasov, Zh. Tekh. Fiz. **72** (9), 1 (2002) [Tech. Phys. **47**, 1071 (2002)].
6. S. K. Demin and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **29**, 1308 (1989).
7. E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Dokl. Akad. Nauk SSSR **314**, 589 (1990) [Sov. Phys. Dokl. **35**, 799 (1990)].
8. R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **31**, 1515 (1992).
9. R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **33**, 1815 (1993).
10. R. P. Tarasov, Radiotekh. Élektron. (Moscow) **44**, 920 (1999).
11. E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **31**, 40 (1992).
12. E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **35**, 1582 (1995).
13. R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **39**, 943 (1999).
14. S. K. Demin and R. P. Tarasov, Mat. Model. **5**, 113 (1993).
15. E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **33**, 1030 (1993).
16. E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **38**, 734 (1998).
17. L. V. Bad’in, Preprint No. 131, IPM AN SSSR (Institute of Applied Mathematics, Moscow, 1982).
18. R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **36** (11), 44 (1996).
19. O. F. Antonenko, *Computing Systems* (IM SO AN SSSR, Novgorod, 1964), No. 12, pp. 39–47 [in Russian].

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 249–258. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 110–120.
Original Russian Text Copyright © 2005 by Bad'in, Zyuzin, Safronov, Slavnov, Tarasov.

ELECTRON AND ION BEAMS, ACCELERATORS

Computer Analysis of the Source Image Structure in 3D Electron–Optic Systems: II. A Stationary Electromagnetic System

L. V. Bad'in, L. N. Zyuzin, S. I. Safronov, Yu. K. Slavnov, and R. P. Tarasov
Research Institute of Pulsed Technology, Moscow, 115304 Russia
e-mail: crabro@aport.ru

Received February 24, 2004; in final form, June 25, 2004

Abstract—Computer analysis of the image of an arbitrary (point or extended) source obtained in 3D electron–optic systems is performed. The systems involve magnetostatic and electrostatic fields, which, respectively, focus and deflect the electron beams. Two approaches to image scanning are considered where the scan potentials are applied in two (symmetric and asymmetric) modes. It is shown that the spread function of the source, which characterizes the imaging quality of the system most fully, may serve as a primary computational criterion. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

The electron–optic systems (EOSs), including the scan units, of pulsed electron-to-optic image converters where imaging is accomplished under the action of the magnetic field produced by a dc coil may basically offer a femtosecond time resolution. An integral stage in designing magnetic (as well as electrostatic [1]) EOSs is computer analysis of them.

A 3D mathematical model of such an EOS is constructed on the assumption that the net electromagnetic field in it is stationary. An axisymmetric magnetostatic field here is produced by a dc coil, and an electrostatic field (including that deflecting the electron beams) is generated by the electrodes under given potentials. A computer model of such EOSs is realized in the form of a program package.

Below, a devised program package is used to analyze the effect of image scan modes (one-phase or asymmetric and two-phase or symmetric) on the resolution of a magnetic-focusing EOS. Image scanning is carried out in two ways: using a pair of separate plates and using a pair of plates with two (entrance and exit) slit diaphragms.

1. BASIC DIAGRAM OF AN ELECTRON–OPTIC SYSTEM

The basic diagram of an EOS with magnetic focusing and electrostatic scanning is depicted in Fig. 1. Here, S_1 is the cathode; S_2, fine-mesh grid; S_3, anode; S_4, deflecting plates; S_5 and S_6, slit diaphragms; S_7, screen; and S_8, dc coil. Conducting electrodes S_i (i = 1–7) are under fixed potentials, S_2 being assumed to be transparent for electrons.

2. COMPUTER MODEL OF AN ELECTRON–OPTIC SYSTEM WITH A STATIONARY ELECTROMAGNETIC FIELD

The electromagnetic field will be assumed to be stationary. Then, the electrostatic and magnetostatic problems can be solved separately.

The electrostatic field is generated by the electrodes kept under given potentials. The statement of this field problem and its numerical solution by the finite-group method (FGM) are considered in [1].

Unlike a potential electrostatic field, a magnetostatic field of strength H is a vortex field. A numerical solution to linear field problems of magnetostatics can generally be obtained by solving boundary integral equations.

A dc coil generates an axisymmetric magnetic field. Such magnetostatic fields can be calculated by analytically extending the field from the axis of symmetry into the interior of the domain. If $H(r, z) = (H_r, 0, H_z)$ is the strength of an axisymmetric magnetostatic field in the cylindrical coordinate system (r, φ, z) and z is the axis of revolution, the magnetic field components can be

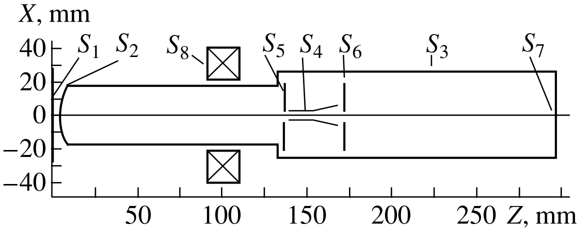


Fig. 1.

1063-7842/05/5002-0249$26.00 © 2005 Pleiades Publishing, Inc.

represented as

$$\begin{aligned} H_z(r,z) &= H_z(0,z) - \frac{r^2}{4}H_z''(0,z) + \dots, \\ H_z(r,z) &= -\frac{r}{2}H_z'(0,z) + \frac{r^3}{16}H_z^{(3)}(0,z) + \dots. \end{aligned} \tag{1}$$

It is assumed that the magnetic field distribution at the axis of the system, $H_z(0, z)$, either is set analytically or is known from experimental data.

Such a technique, being rather simple, has found wide application. At the same time, the procedure of analytical extension is conditionally stable in numerical terms. Therefore, an algorithm implementing relationships (1) becomes of crucial importance.

The calculation scheme for analytical extension that is embodied in the program package for EOS simulation allowed us to construct a magnetic field in the cylindrical domain $0 \le r \le (1/2)R$, where R is the radius of the anode chamber, with a high accuracy.

The characteristics of the system, as in the case of electrostatic EOSs, are found by means of trajectory analysis [1].

Electron trajectories in a stationary electromagnetic field are described by the Cauchy problem for the set of partial differential equations

$$\frac{d}{dt}\left(m\frac{dr}{dt}(x,t)\right) = -e\nabla\varphi(x) + \frac{e}{c}\left[\frac{dr}{dt}(x,t)H(x)\right], \tag{2}$$

where r is the radius vector of a particle with mass m and charge e and c is the speed of light (for the initial conditions of particle escape, see [1]).

In relativistic case (2), numerical integration of the equations of motion is performed by standard (implicit or explicit) second-order schemes involved in the finite-difference method. It is essential that here, as in the case of electrostatic EOSs, adaptive beam-attendant spatial computational grids are used.

As before [1], the image of point emitter P on given surface s will be characterized by the center of gravity τ of the electron beam emitted, rms deviation σ of the beam from center of gravity τ, and spread function $S_p(i, j)$ of the point emitter.

The calculation results will be visualized in Cartesian coordinate system XYZ where the origin is placed at the center of the cathode and the Z axis is aligned with the optical axis of the EOS; cylindrical coordinate system r, ψ, z related to the XYZ system; and Cartesian system $X'Y'Z'$ to visualize the spread functions. Coordinate system $X'Y'Z'$ is obtained from XYZ by rotation through the angle equal to the angle of rotation of the image on the screen. In this case, the electrons emitted from cathode points $(0, y, 0)$ are collected on the screen along the Y' axis provided that the plates are grounded.

3. IMAGE STRUCTURE IN AN AXISYMMETRIC ELECTRON–OPTIC SYSTEM

Unlike an electrostatic field, a magnetostatic field imparts not only translational but also rotational motion to electrons. This necessitates a close look at the electron gun of an EOS (an axisymmetric EOS without scanning is meant, Fig. 1).

Imaging by means of the electron gun may be viewed as bringing electrons emitted by point emitters on the cathode surface to a net beam that has a minimal cross-sectional area (crossover) in a certain plane $Z = Z_c$ (the plane of crossover). In axisymmetric focusing EOSs, the position of the crossover and the net beam structure in the plane of crossover virtually completely define the imaging quality, since, in these EOSs, the electron trajectories downstream of the plane of crossover are close to straight lines. Therefore, computer analysis of net beam passage through the EOS working region (from the cathode to the screen) and approximation of the subsequent trajectories by straight lines give a comprehensive estimate of the performance of the electron gun. Such an approach is efficient in determining the shape and position of the image surface; image structure on the image surface, screen, or any secant plane of interest; angle of rotation of the image, magnification on a given plane; position of the crossover and the structure of the net beam in it; and position of the scan unit. Moreover, one may also extract the time characteristics with this approach (they are omitted from the present consideration).

Below, we report computer analysis data on the behavior of a net beam consisting of 11 elementary beams in the working region of the electron gun (Fig. 1) (the scan unit is under the potential of the anode chamber, i.e., grounded). Each of the elementary beams that is emitted from a cathode point (r, φ = const, r = 0, 0.5, 1.0, …, 5.0 mm) may trace 512 equiprobable trajectories on assumption that the probability of a particle escaping within a solid angle $d\omega$ at an angle θ to normal n_p to point P varies as $\cos^3\theta d\omega$ [1]. As follows from computing experiments, this number of trajectories suffices to study the structure of an elementary electron beam in the EOS under consideration. We also assume that the electrons have the same escape energy U_0; that is, the energy spread in the source is disregarded. The results that follow were obtained for U_0 = 4 eV and a potential difference between the cathode (S_1) and grid (S_2) of 10 kV.

(i) The parameters of the magnetic lens were selected such that the beam emitted from the cathode point with the radial coordinate $r = r_7 = 3$ mm is focused on the screen (i.e., the rms deviation σ_7 from the center of gravity τ_7 of the beam is minimal). When focused, the centers of gravity of the 11 beams lie on the generatrix of a surface of revolution that is the image surface. In this case, σ_i (i = 1, 2, …, 11) take values from the segment [0.01, 0.15 mm] and the image surface is fairly

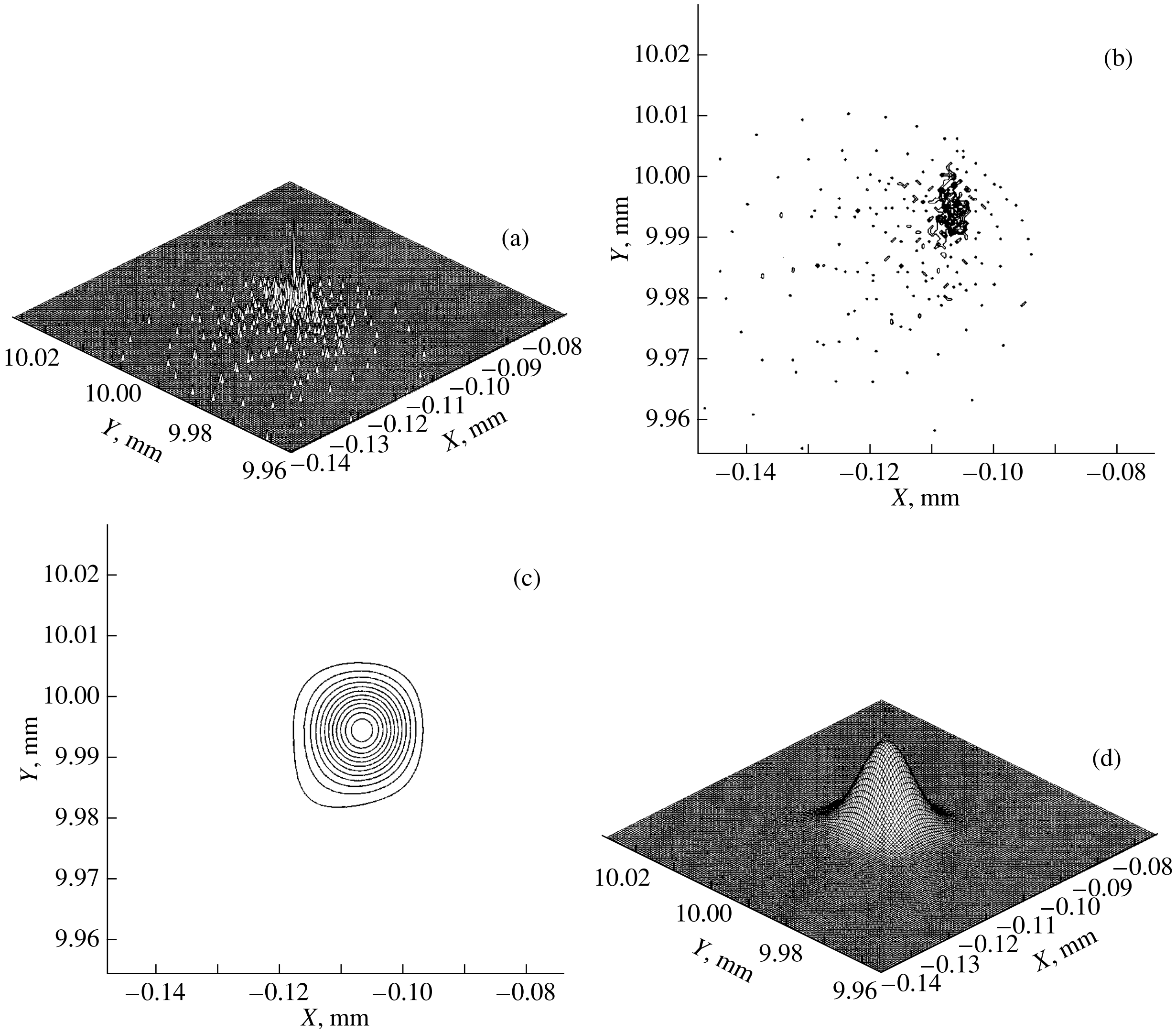


Fig. 2.

accurately approximated by part of an ellipsoid of revolution.

Figures 2a and 2b show, respectively, the values and isolines of the spread functions $S_{r_{11}}(i, j)$ of the point emitter with the radial coordinate $r_{11} = 5$ mm (the edge of the cathode) on the image surface. When visualized on the image surface and *a fortiori* on the screen, the spread functions of the elementary beams look much different. It should be emphasized, however, that discrete spread functions are hard to compare visually: an extra smoothing procedure is needed.

Spread functions $S_{r_k}(i, j)$, $k = 1–11$, were smoothed with the same *B*-spline digital filter [1]. Smoothed spread functions $S_{r_k}(i, j)$ will be designated as $\hat{S}_{r_k}(i, j)$, $k = 1–11$.

On the image surface, functions $\hat{S}_{r_k}(i, j)$ have a well-defined bell-shaped form, which somewhat deforms with distance from the axis of the system. Figures 2c and 2d demonstrate, respectively, the isolines and values of spread function $\hat{S}_{r_k}(i, j)$. On the screen, function $\hat{S}_{r_7}(i, j)$ alone has a bell-shaped form. This is because the beams emitted from cathode points r_i, $i = 1–6$, are focused behind the screen plane and those emitted from points r_i, $i = 8–11$, are focused before the screen. Accordingly, functions $\hat{S}_{r_i}(i, j)$, $i = 1–6$, on the screen have the form of a flattened bell and bell-shaped functions $\hat{S}_{r_i}(i, j)$, $i = 8–11$, have a clearly cut substrate (diffuse base).

The advanced digital processing techniques make it possible to comprehensively analyze the image structure in an EOS by computer methods. For instance,

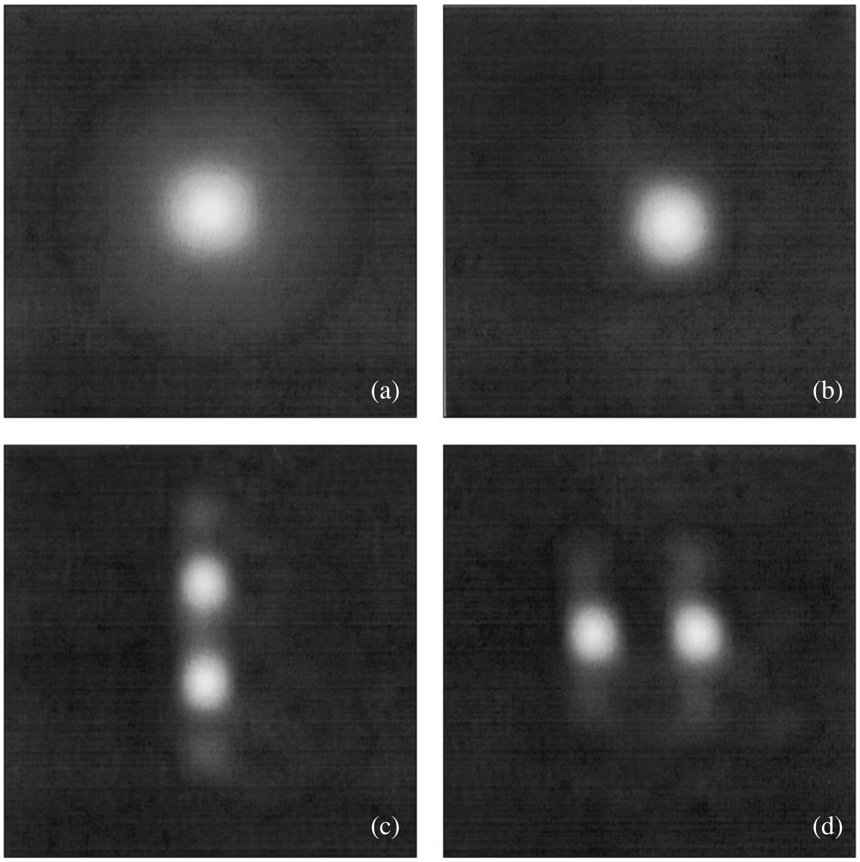


Fig. 3.

Fig. 3a shows the "photograph" of the point emitter image ($r_1 = 0$, the center of the cathode) on the screen with a 1000-fold magnification (the photograph was obtained by digital processing methods); Fig. 3b, the photograph of the point emitter with $r_{11} = 5$ mm; and Figs. 3c and 3d, the photographs of the emitter r_{11} and the emitter 0.01 mm distant. In Fig. 3c, the emitters lie on the same ray passing through the center of the cathode; in Fig. 3d, the emitters lie on a normal to this ray. In both cases (Figs. 3c and 3d), the images on the screen are separated, the separation being more distinct in Fig. 3d.

(ii) Computer analysis of mathematical models that are developed for technical systems (EOS in our case) is of practical value if computation is carried out with split-hair accuracy. It is, as a rule, assumed that the mathematical model of a given system adequately describes physical processes in it. It would be good to compare the system's characteristics found experimentally and analytically; however, this is a challenge sometimes for a number of reasons (relevant experimental data are impossible to obtain, a high cost of the experiment, rough measurements, etc.). At the same time, computational tests, representative as they may be, cannot serve as an undeniable indication of validity of computer analysis. Moreover, a known test problem that is fully adequate to a given mathematical model usually can be selected only in simplest cases. Fortunately, there is a test problem that is applicable (to a certain degree of confidence) to the EOS 3D mathematical model under consideration.

Gaussian dioptrics assumes that systems with an axisymmetric stationary electromagnetic field have a cylindrical region G of small radius $r = R_G \ll L$ (L is the EOS length) that encircles the axis of revolution of the system and where the radial components of the electromagnetic field linearly depend on r. If so, paraxial monochromatic beams tracing smooth trajectories that entirely fall into axial region G are focused into a point. If, in addition, there exists an external electromagnetic field that deflects the electron beams and weakly interacts with the starting focusing field, the paraxial beam is deflected as a whole, i.e., again is focused into a point.

The EOS mathematical model considered in this work is stated in a much more general manner than Gaussian dioptrics. At the same time, if we assume that cylindrical region G of radius R_G is accessible to machine calculation ($R_G/L > 10^{-5}$) and that the computer model inaccuracy (associated with the need for solving axisymmetric and 3D field problems and integrating the equations of motion) is comparable to R_G/L, there appears a possibility of Gaussian dioptrics being numerically implemented in the framework of the EOS computer model mentioned above. To do this requires that solid angle $d\omega$ to which the elementary paraxial beams are confined (180° in the previous calculations) be considerably decreased.

Figure 4a shows the isolines of spread function $S_{r_1}(i, j)$ for the central electron beam ($r_1 = 0$) with solid opening angle $d\omega = 180°$ on the optimal image surface. Here, the rms deviation of this beam, $\sigma_1(d\omega) = \sigma_1(180°) = 0.01$ mm, is almost equal to the radius of the spot of confusion. Having fixed all the basic parameters of the elementary beam under consideration (the number of trajectories (512), the electron escape energy, and the electron angular distribution), we will see how the size of the spot of confusion varies with the opening angle $d\omega$ of the beam.

The threefold decrease in the opening angle ($d\omega = 60°$) does not cause any appreciable change in the spot of confusion (Fig. 4b), and the characteristic size of this spot for the electron beam with $d\omega = 20°$ (Fig. 4c) is only ≈2.2 times smaller than its initial value (for $d\omega = 180°$). At the same time, the central beam with $d\omega = 10°$ is focused to a "point of calculation" (Fig. 4d); that is, the characteristic size of the spot of confusion on the image surface becomes 200 times smaller than the initial value, $\sigma_1(10°) = 5 \times 10^{-5}$ mm.

A shift of the point emitter from the cathode center ($d\omega = 10°$) results in a rapid diffusion of the focus: the beam emitted from the cathode point with the radial coordinate $r = 0.5$ mm has a spot of confusion that is comparable in size to the spot of the initial central beam with $d\omega = 180°$ (Fig. 4e). Then, when the scan potentials applied to the deflecting plates equal ±30 V (the slit diaphragms are under the potential of the anode chamber, Fig. 1), the central paraxial beam ($r_1 = 0$, $d\omega = 10°$) is deflected by 3 mm virtually as a whole ($\sigma = 8 \times 10^{-5}$ mm). At the same time, as the scan potential rises to ±100 V, the spot of confusion of the central paraxial beam on the image surface substantially expands to $\sigma \approx 0.004$ mm (Fig. 4f).

Thus, the basic concepts of Gaussian dioptrics are implemented in terms of our computer model of EOSs with an electromagnetic field.

(iii) Investigation into the behavior of a bunch of elementary beams as a whole (net beam) in the EOS working region makes it possible to find the position of the crossover and determine the net beam structure in it. Other parameters that can be determined are the position of the deflecting plates and an optimal angle between them that eliminate vignetting, as well as the angle through which the image should be rotated. To illustrate the aforesaid, let us consider a bunch of six elementary beams emitted from cathode points (r, ψ = const), $r = 0, 1, \ldots, 5$ mm. Each of the beams is assumed to cover 32 equiprobable trajectories.

Figure 5 visualizes the structure of the net beam in different cross sections ($z = 0$). Figure 5a refers to the section where the effect of the magnetic field (rotational moment) on the net beam becomes tangible. Figure 5b shows the net beam structure in the plane of crossover. Figures 5c and 5d illustrate the situations when the net beam, respectively, comes into and goes out of the space between the deflecting plates. Figure 5e demonstrates the net beam structure in a section between the scan unit and screen. Finally, Fig. 5f visualizes the net beam structure on the screen. Note the angle of rotation of the image remains nearly constant from the entrance to the deflecting space to the screen. The transverse sizes of the net beam at the entrance to and the exit from the deflecting space specify a minimal distance between the deflecting plates that excludes vignetting with the scan unit grounded.

4. EFFECT OF ELECTROSTATIC SCANNING MODES ON THE IMAGE STRUCTURE

Let us set the task of elucidating the influence of the one-phase (asymmetric) and two-phase (symmetric) scan modes on the image structure in an EOS. Two types of scan units are considered: a pair of separate deflecting plates and a pair of plates with slit diaphragms (Fig. 1). In the one-phase scan mode, one plate is grounded and the other is under a given potential. In the two-phase mode, the deflecting potentials are equal in magnitude but opposite in sign.

(i) Separate deflecting planes. The EOS resolution versus scan potential φ for the symmetric and asymmetric scan modes was studied as follows. Given φ, six elementary beams escaping the cathode ($z = 0$) with escape coordinates ($x = 0$, y_i; $i = 0, 1, \ldots, 5$) were calculated ($y_0 = 0$, $y_1 = 1$, …, $y_5 = 5$ mm). Each of the beams consisted of 512 equiprobable trajectories (the electron angular distribution for $U_0 = 4$ eV is given above). Then, the least rms deviation $\sigma_\varphi(i)$ ($i = 0, 1, \ldots, 5$) was calculated for each of the beams; that is, the focus of the beams was found in terms of this criterion. The values of $\sigma_0(i)$ (namely, $\sigma_0(0) = 0.009$ mm, $\sigma_0(1) = 0.009$ mm, $\sigma_0(2) = 0.015$ mm, $\sigma_0(3) = 0.013$ mm, $\sigma_0(4) = 0.014$ mm, and $\sigma_0(5) = 0.015$ mm) correspond to the case of the grounded plates ($\varphi = \pm 0$ V, axisymmetric system), with the beam originating from point ($x = 0$, $y_3 = 3$ mm) being focused on the screen.

For the symmetric scan mode, Table 1 lists ratios $\sigma_\varphi(i)/\sigma_0(i)$ ($i = 0, 1, \ldots, 5$) for scan potentials $\varphi = \pm 90$, ± 180, ∓ 90, and ∓ 180 V.

From Table 1, one can see that the resolution on the image surface may drop 1.33–5.10 times depending on scan potential φ and the point emitter position. Column X shows the values of coordinate x (the scan direction) of the spread centers for all the six elementary beams on the screen.

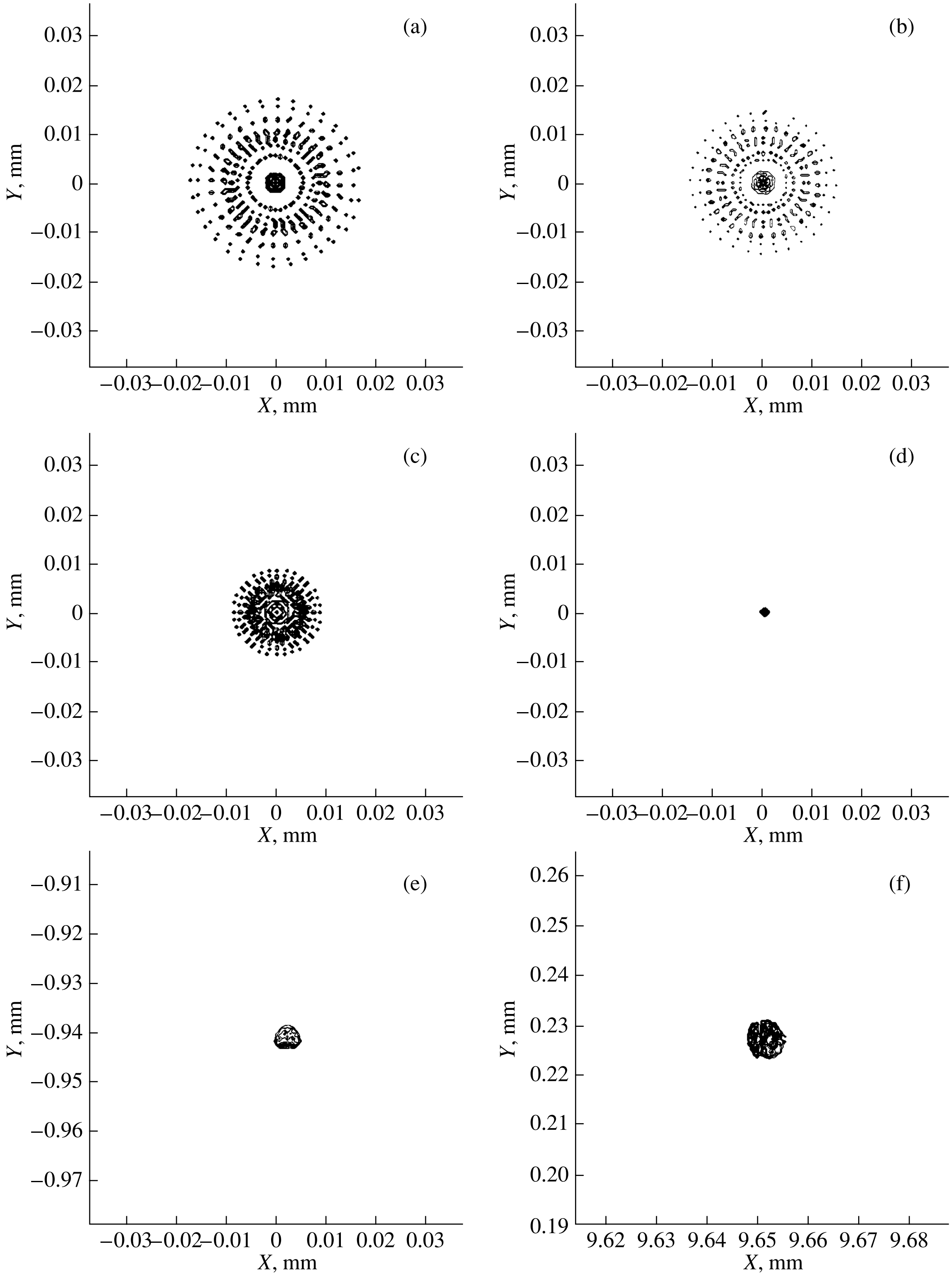


Fig. 4.

For the asymmetric scan mode, ratios $\sigma_\varphi(i)/\sigma_0(i)$ ($i = 0, 1, \ldots, 5$) for scan potentials $\varphi = 180, 360, -180$, and –360 V are given in Table 2. In this case, the resolution of the instrument on the image surface may drop 6–24 times depending on the scan potential and the point emitter position.

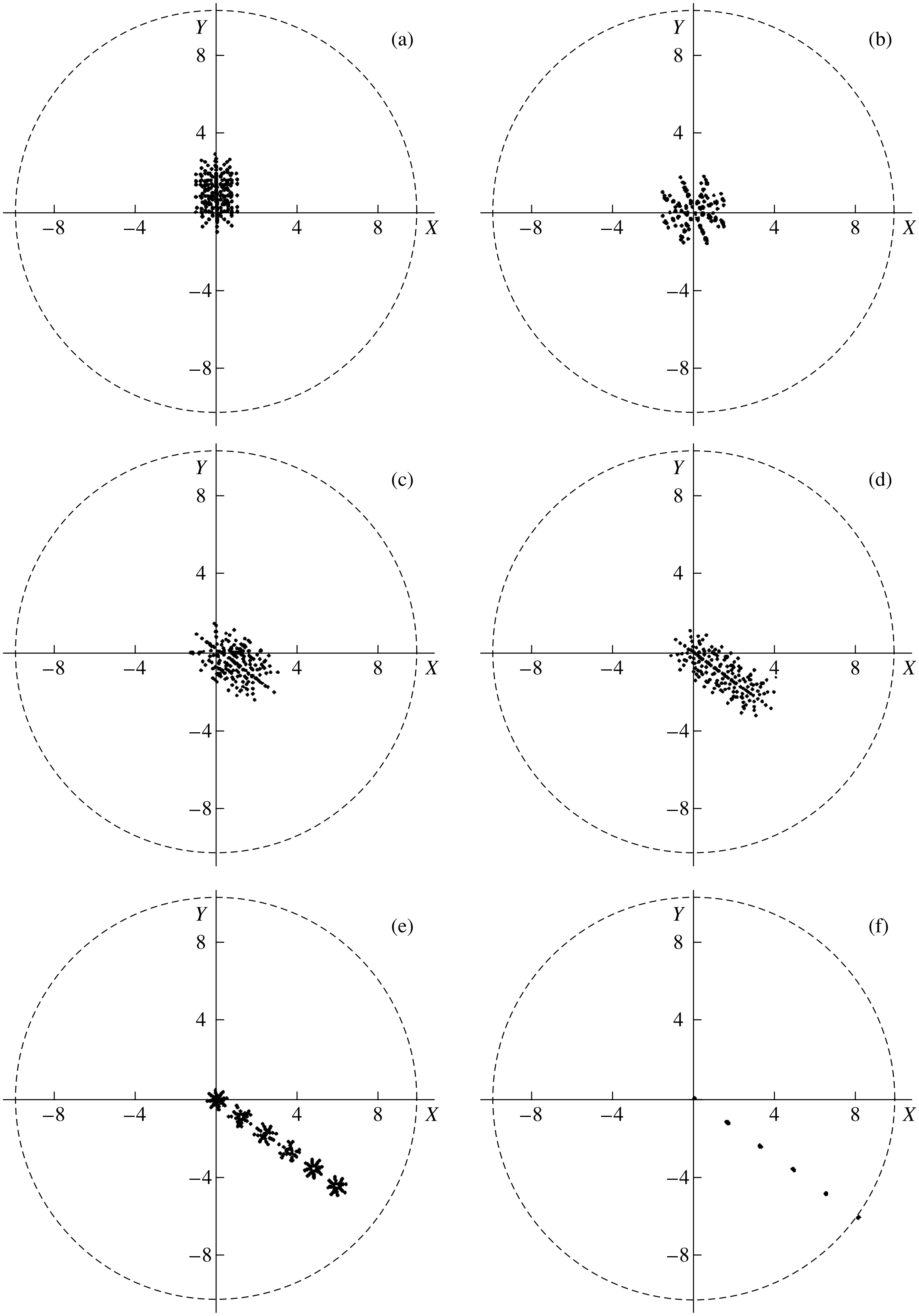


Fig. 5.

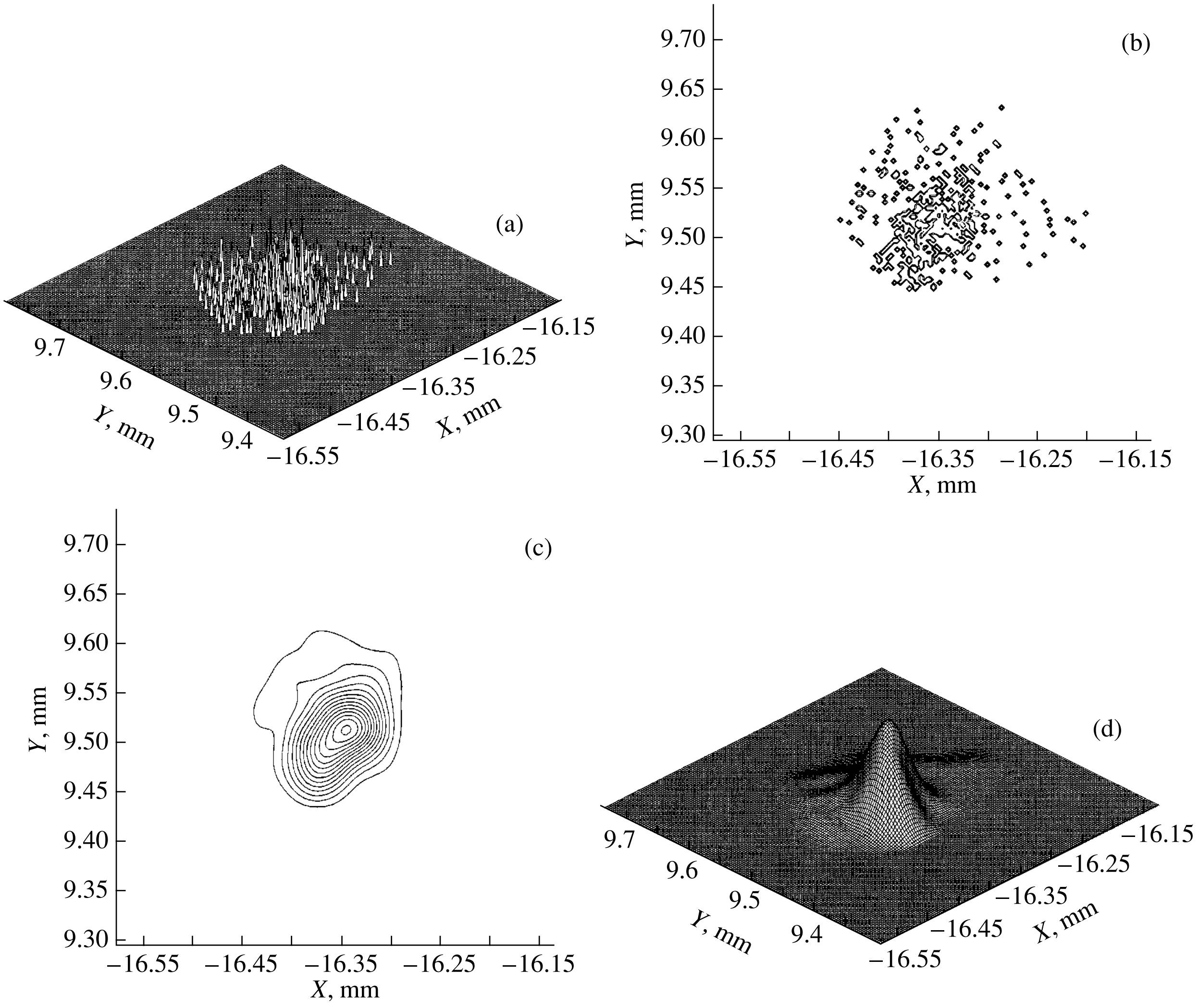


Fig. 6.

In the symmetric and asymmetric scan modes, the image surfaces do not coincide and share a single line of intersection, which is the generatrix of the image surface in an axisymmetric EOS. Also, the deflection of the elementary beam is virtually completely defined by potential difference Δφ between the deflecting plates (columns X in Tables 1, 2). Consequently, the resolutions of the instrument on the image surfaces that are observed in the two scan modes can be well compared if Δφ is fixed. Ratio $\sigma^1_{\Delta\varphi}(i)/\sigma^2_{\Delta\varphi}(i)$ (where $\sigma^1_{\Delta\varphi}(i)$ and $\sigma^2_{\Delta\varphi}(i)$ are the rms deviations in the one- and two-phase scan modes, respectively) shows in which of the modes and how much the resolution on the image surface is higher.

Table 1. Two-phase scan mode

φ, V\\y_i, mm	0	1	2	3	4	5	X, mm
±180	5.1	5.1	3.33	4.08	4.14	4.6	17.2
±90	1.78	1.89	1.67	2.15	2.36	2.8	8.65
±90	1.78	1.67	1.33	1.77	2.14	3	–8.65
∓180	5.1	4.89	3.2	3.85	3.79	4.93	–17.2

The values of $\sigma^1_{\Delta\varphi}(i)/\sigma^2_{\Delta\varphi}(i)$ ($i = 0, 1, \ldots, 5$) for $\Delta\varphi =$ 180, 360, –180, and –360 V are given in Tables 1 and 2. As follows from these tables, the resolution on the image surface is two to seven times higher under the two-phase scan conditions. Moreover, the depth of the image surface along the z coordinate is 13 mm in the symmetric mode and 30 mm in the asymmetric one. Thus, in the system under consideration, the two-phase scan mode is several times more efficient in terms of imaging.

(ii) Deflecting plates with slit diaphragms. The EOS design (Fig. 1) the scan unit of which consists of deflecting plates with two, entrance and exit, slit dia-

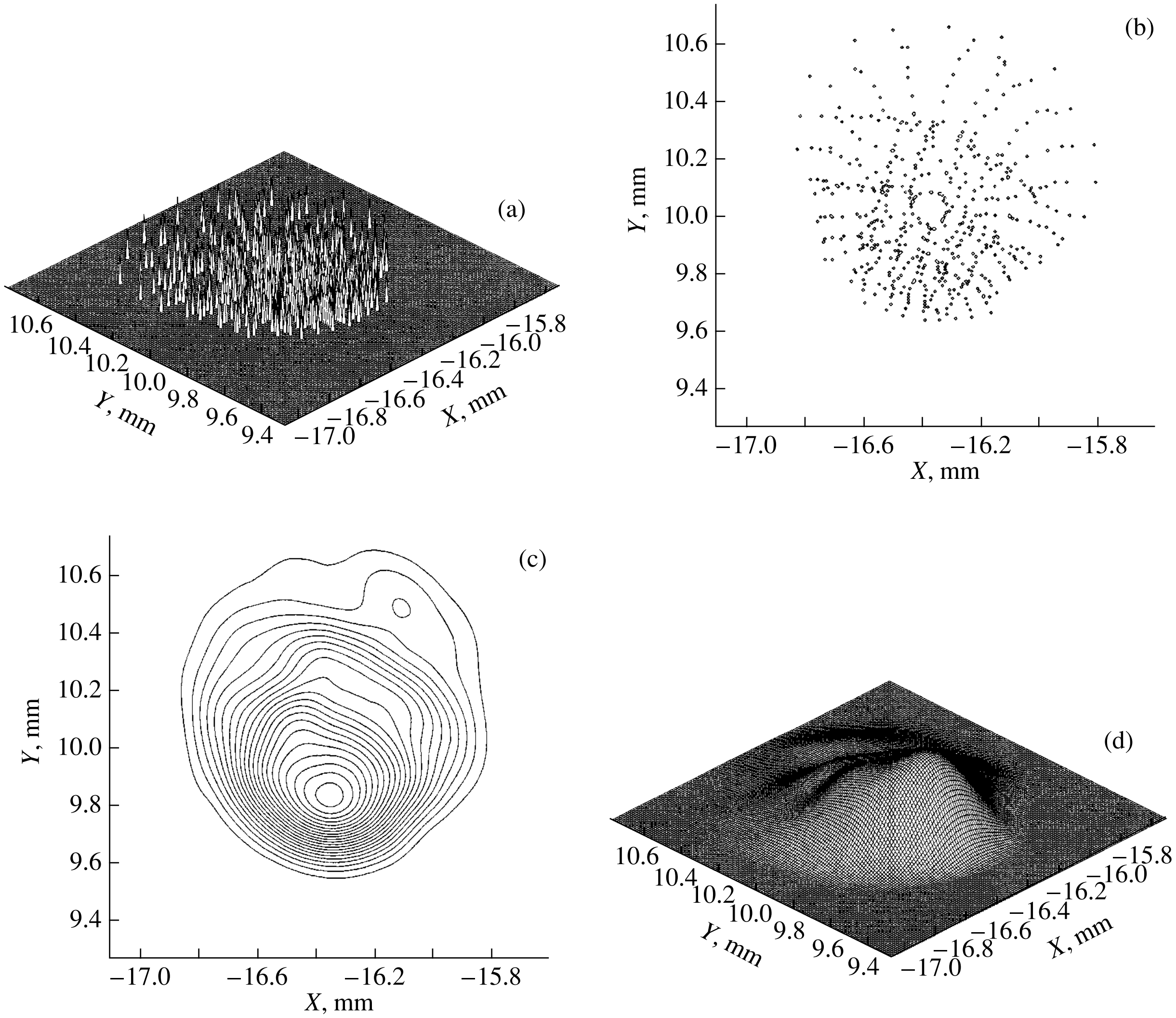


Fig. 7.

phragms will be numerically analyzed following the scheme used to analyze the design with a pair of separated scanning plates (see the previous section).

For the symmetric scan mode and scan potentials $\varphi = \pm 90$ and ± 180 V, ratios $\sigma_\varphi(i)/\sigma_0(i)$ ($i = 0, 1, \ldots, 5$) are listed in Table 3.

The results of calculation for $\varphi = \pm 90$ and ± 180 V are virtually identical to those for $\varphi = \mp 90$ and ∓ 180 V up to sign (Table 3, column *X*). Thus, the resolution of the EOS design under consideration decreases two to seven times on the image surface according to the scan potential and point emitter position.

For the asymmetric scan mode and scan potentials $\varphi = 180, 360, -180$, and -360 V, ratios $\sigma_\varphi(i)/\sigma_0(i)$ ($i = 0, 1, \ldots, 5$) are listed in Table 4.

Under these scan conditions, the resolution on the image surface drops by a factor of 5–18.

Comparing Tables 3 and 4 shows that, for the EOS design considered, the resolution in the two-phase scan mode may exceed that in the one-phase mode by a factor of four at most. Thus, diaphragming somewhat reduces the difference between the two scan modes. Because of this, it would be of interest to estimate the effect of diaphragming on the resolution in either scan mode.

It follows from Tables 1 and 3 that diaphragming in the two-phase mode leads to an insignificant increase (≈10%) in the resolution at the edge of the image surface; in the remaining part of the image surface, the resolution, however, drops. In the one-phase mode, conversely, diaphragming improves the resolution by 30–60% throughout the image field (Tables 2, 4).

The fact that diaphragming variously affects the resolution in the one- and two-phase scan modes is just the reason why the resolutions approach each other in the scan mode with diaphragming. At the same time, the

Table 2. One-phase scan mode

φ, V\\y_i, mm	0	1	2	3	4	5	X, mm
$\begin{pmatrix}360\\0\end{pmatrix}$	14.89	14.22	9.67	12.54	14.71	16.93	17.2
$\begin{pmatrix}180\\0\end{pmatrix}$	9.44	9.56	6.13	8	8.79	10.27	8.65
$\begin{pmatrix}-180\\0\end{pmatrix}$	12	12.11	7.66	9.77	10.64	11.73	–8.65
$\begin{pmatrix}-360\\0\end{pmatrix}$	24.11	23.67	15.06	18.85	19.93	21.47	–17.2

Table 3. Two-phase scan mode

φ, V\\y_i, mm	0	1	2	3	4	5	X, mm
±180	7.11	6.44	4.06	4.38	3.85	4.2	16.3
±90	4	2.33	2.4	2.46	2.21	2.6	8.2

Table 4. One-phase scan mode

φ, V\\y_i, mm	0	1	2	3	4	5	X, mm
$\begin{pmatrix}360\\0\end{pmatrix}$	6.56	6.22	4.93	7.07	10.35	14	16.3
$\begin{pmatrix}180\\0\end{pmatrix}$	5	5	3.73	5.46	6.57	8.3	8.2
$\begin{pmatrix}-180\\0\end{pmatrix}$	8.11	8.11	5.33	7.31	8.36	9.87	–8.2
$\begin{pmatrix}-360\\0\end{pmatrix}$	17.33	17.1	12.85	14.77	16.43	18.53	–16.3

depth of the image surface along the z coordinate is 11 and 27 mm for the symmetric and asymmetric scan modes, respectively. Hence, in this case, too, the two-phase mode seems much more preferable than the one-phase regime. Such a conclusion is totally consistent with the result for electrostatic EOSs [1].

(iii) Structure of the spot of confusion for an elementary electron beam. The σ criterion (σ is the rms deviation of the trajectories from the center of gravity τ of a beam) is a straightforward [1], reliable, and more or less information-carrying criterion in simulating EOSs with an electrostatic and stationary electromagnetic field, in which the spot of confusion is near-circular. Furthermore, this criterion is quite convenient for comparative analysis of basically different operating modes (designs), e.g., of those considered above (a several-fold difference in resolution). At the same time, scanning of the image not merely deforms but considerably reconfigures the spots of confusion of point emitters; consequently, it is necessary to replace the rms deviation by a more informative criterion in studying the image structure. The point-spread functions before, $S_p(i, j)$, and after, $\hat{S}_p(i, j)$, the smoothing procedure characterize the structure of the spot of confusion of an electron beam most comprehensively [1].

In real EOSs, numerical image analysis requires that a great number of spread functions (for different emitters in order to gain a representative set of scan potentials φ) to be known. Here, we will consider only the spread function on the image surface for point emitter P ($x = 0$, $y = -5$, $z = 0$) at scan potentials $\varphi = \pm 0, \pm 180$, and –360 V (asymmetric scan mode) in the case of slit diaphragm scanning (Fig. 1).

The values and isolines of the spread functions before, $S_p^{(1)}(i, j)$, and after, $\hat{S}_p^{(1)}(i, j)$, filtering at $\varphi = \pm 0$ V are given in Fig. 2. For the two-phase scan mode at $\varphi = \pm 180$ V, the values and isolines of $S_p^{(2)}(i, j)$ are given in Figs. 6a and 6b, respectively. Figures 6c and 6d show the isolines and values of $\hat{S}_p^{(2)}(i, j)$, respectively. Similar information on spread functions $S_p^{(3)}(i, j)$ and $\hat{S}_p^{(3)}(i, j)$ in the asymmetric scan mode ($\varphi = -360$ V) is visualized in Fig. 7.

Spread functions $S_p^{(i)}(i, j)$ ($i = 1$–3; Figs. 2a, 6a, 7a) and especially their isolines (Figs. 2b, 6b, 7b) clearly demonstrate that the rms deviation as a criterion for studying the image structure is inappropriate. At the same time, the isolines (Figs. 2c, 6c, 7c) and values (Figs. 2d, 6d, 7d) of spread lines $\hat{S}_p^{(i)}(i, j)$ ($i = 1$–3) corroborate the validity of the scan mode analysis performed above.

Note in conclusion that use of point-spread functions greatly extends the potentialities of numerical analysis when studying the image structure in various EOS designs (Fig. 3) and makes it possible to study not only stationary spatial characteristics but many of others as well.

REFERENCES

1. L. V. Bad'in, L. N. Zyuzin, S. I. Safronov, *et al.*, Zh. Tekh. Fiz. **75** (2), 99 (2005) [Tech. Phys. **50**, 237 (2005)].

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 259–261. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 121–123.
Original Russian Text Copyright © 2005 by Matviishyn, S. Novosad, I. Novosad.

BRIEF COMMUNICATIONS

Influence of Adsorption on the Transverse Photovoltage Arising in Cadmium Diiodide Crystals under X-ray Irradiation

I. M. Matviishyn, S. S. Novosad, and I. S. Novosad

Franko State University, ul. Universitetskaya 1, Lviv, 79000 Ukraine

Received January 8, 2004; in final form, May 21, 2004

Abstract—The influence of a gas atmosphere (O_2 or air) on the transverse photovoltage (TPV) induced by X-ray irradiation in CdI_2 specimens is studied. It is shown that the TPV measured at 295 K in a vacuum grows almost linearly with X-ray dose rate R. In a gas atmosphere, the TPV is of opposite sign and tends to saturation when R exceeds 200 R/min. After evacuation of the system, the TPV value increases, attains a maximum, and then abruptly drops (changing sign) to the initial "vacuum" value. Based on the results of this study and published data, it is shown that such behavior of the TPV in CdI_2 is related to X-ray-stimulated chemisorption proceeding by the acceptor mechanism. © *2005 Pleiades Publishing, Inc.*

The processes induced by high-energy radiation at the solid–air interface are today among the most topical and yet poorly studied problems in radiation materials science [1]. Interest in these processes is also heightening because of the search for promising materials for high-sensitive gas analyzers [2].

The transverse photovoltaic effects arising in layered CdI_2 crystals exposed to optical and X-ray radiation in a vacuum were considered in [3–5]. In this study, we investigate the influence of gases (O_2 and air) on the transverse photovoltage arising in X-ray-irradiated CdI_2 crystals.

Specimens from CdI_2 crystals grown by the Bridgman–Stockbarger technique [6] were prepared and studied in the same way as described in [3, 4]. The photovoltage was measured in the open-circuit mode, U_{oc}, on the specimens prepared in the form of ~10 × 8 × (0.5–6.0)-mm parallelepipeds such that the normal to the surface irradiated (10 × 8 mm) made an angle of 0° and 45° with crystallographic axis C_6. Ohmic contacts made of K-13b silver paste were deposited on the 8 × (0.5–6.0)-mm lateral surfaces. The specimens with the contacts were placed into a metallic cryostat, where the measurements were performed at 295 K both in a vacuum and in the gas media. URS-55A equipment with an VSV-2 copper-anode X-ray tube (U = 45 kV, I = 0–12 mA) served as an X-ray source.

In the 0°-cut specimens measured in a vacuum, the TPV value is almost independent of their thickness. For both the 0°- and 45°-cut specimens measuring 10 × 8 × 6 mm, the TPV grows linearly with dose rate R. Curve *1* in Fig. 1 shows this dependence for the 0°-cut specimen. The 45°-cut specimen is approximately 2.5 times more sensitive to X rays. The TPV values measured on the layered cadmium iodide crystals in a vacuum agree with the data reported in [3–5, 7].

X-ray irradiation of CdI_2 crystals at atmospheric pressure in air also generates a TPV. However, in this case, the voltage is of opposite sign and tends to saturation at R > 200 R/min (Fig. 1, curve *2*). For the CdI_2 specimens in the O_2 atmosphere, the dependence is similar but the saturated value of the TPV is 20–25% higher. For the irradiation in air, the TPV induced in the 0°-cut specimen 0.5 mm thick is nearly half that measured in the 1.5-mm-thick specimen.

The dependence of the TPV generated in the 45°-cut CdI_2 crystal on the oxygen pressure at R = 700 R/min is shown in Fig. 2. As the pressure decreases, the TPV grows, peaks, and then abruptly drops, changing sign and saturating at the initial (vacuum) level. In air, the

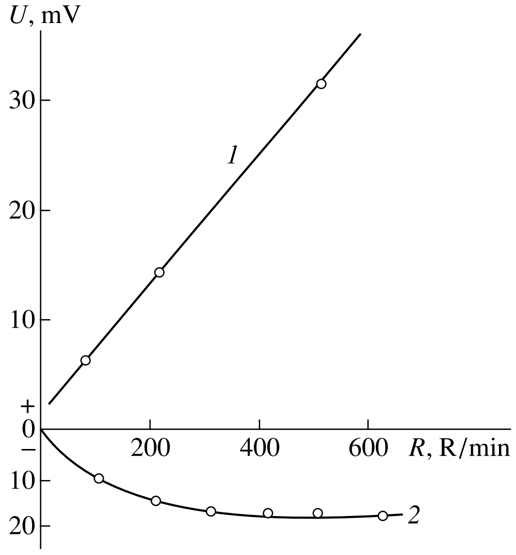


Fig. 1. TPV vs. R for the 0°-cut CdI_2 crystal (*1*) in a vacuum and (*2*) in air.

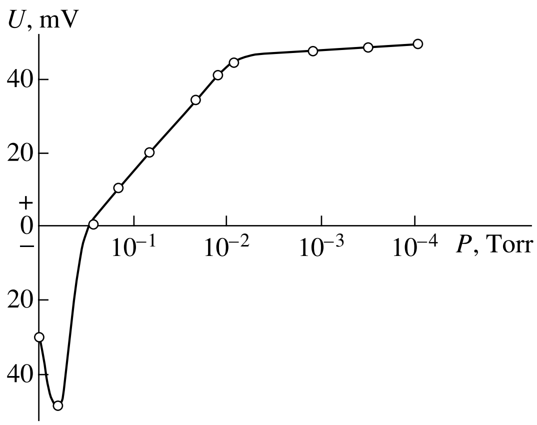


Fig. 2. TPV vs. the oxygen pressure for the 45°-cut CdI_2 crystal at R = 700 R/min.

run of the TPV versus pressure dependence is the same. Such behavior is also typical of the 0°-CdI_2 crystals.

When O_2 is first introduced into the cryostat during the irradiation and then is pumped off, the TPV exhibits a spike. In addition, the voltage changes sign when the gas is both introduced and pumped out. As compared to the curve in Fig. 2, such a dependence shows an extra peak at the instant of gas delivery.

It was found that the TPV resulting upon the X-ray irradiation in oxygen strongly (nearly exponentially) varies with temperature in the interval from 295 to 380 K. In the high-temperature range, the dependence $U_{oc} = f(R)$ becomes even linear at $R > 150$ R/min. With a decrease in the crystal temperature to 295 K, the TPV takes its initial value. Preirradiation of the specimens in a vacuum does not affect their adsorptivity with respect to air and oxygen either.

The results presented above suggest that a gas atmosphere has a considerable effect on the value of a TPV induced by X rays in CdI_2 and may change its sign. It was shown [3–5] that transverse photovoltaic effects in layered CdI_2 crystals exposed to optical or X-ray radiation in a vacuum appear as the photogalvanic effect and the transverse Dember effect [8], since the structure, chemical bonds, and electric conductivity in these crystals are highly anisotropic. The saturation of the TPV in the gas media at 295 K and $R > 200$ R/min may be attributed to competition between adsorption and the process associated with the Dember mechanism [7], with adsorption prevailing at high temperatures.

An intriguing feature of our observations is that the R dependence of the TPV in CdI_2 at 295 K under atmospheric pressure (Fig. 1, curve *2*) is similar to that taken in a vacuum at 90 K [8]. At low temperatures, cadmium diiodide has *n*-type conductivity, while at 295 K, unexcited cadmium iodide is a *p*-type semiconductor. Irradiation by X rays at this temperature generates donor centers at the surface and the surface takes a negative charge [8]. Sorption of acceptor-type gases (O_2, NO_2, N_2O, or CO_2) [9, 10] reduces the electron concentration at the surface and lowers the conductivity of the specimen. Simultaneously, the concentration of holes near the surface grows. The competition between the sorption on the surface and the generation of nonequilibrium charge carriers in the bulk specifies the sign and absolute value of the TPV as a function of pressure in the system [10].

Since the TPV takes the initial value and sign after the gas has been pumped off, one can suppose that X-ray-stimulated adsorption takes place on the surface irradiated. The charge carriers trapped by surface defects are likely to act as absorption centers, short-lived defects playing a major part in the sorption process [10].

The participation of free charge carriers in the adsorption-related processes on the surface, as well as the fact that these processes are temperature-activated, suggest that the gases influence the transverse photoeffects via chemisorption [10–13]. It is implied that gas particles are adsorbed on the surface, capturing free electrons, rather than penetrate into the crystal [14].

The chemisorption produces a local surface energy level, which is similar to normal impurity levels related, for example, to surface defects [12, 13]. Localization of charge carriers on surface electron states charges the surface. If free carriers are present in the bulk, they are redistributed so as to neutralize the surface charge. A double electrical layer thus forms near the surface, with its thickness depending on the concentration of free carriers and ionized impurities.

Thus, the efficiency of chemisorption to a large extent depends on the charge concentration near the surface and on the density of surface states. Under stationary conditions, adsorption on the semiconductor surface comes into equilibrium when two processes equilibrate with each other: (i) exchange between particles adsorbed and the gas medium and (ii) electron exchange between the electron–hole system of the crystal and the adsorption-produced level [10–12]. In this system, surface and bulk charges generate an intense electric field directed normally to the surface irradiated. Under sorption, electron (hole)-enriched or depleted regions may be as deep as 0.5–50 nm [12].

The thinner the 0°-cut crystal, the more intense the chemisorption on the rear surface due to X rays that extend through the crystal. This generates an extra electric field opposing the field caused by the gas adsorption on the front surface. Competition between the fields diminishes the TPV value. The increase in the TPV early in the pumping may be due to the fact that the absorption of X rays by the gas medium and charge exchange between the particles adsorbed and the gas weaken. The sharp fall of the TPV and reversal of its sign during pumping are related to a decrease in the number of particles adsorbed on the surface.

REFERENCES

1. A. E. Toropov, I. A. Vasilev, and A. F. Nechaev, *Radiation Induced Phenomena in Solid: Institute Collection of Scientific Works* (UPI im. S. M. Kirova, Sverdlovsk, 1979), No. 1, pp. 54–61 [in Russian].
2. V. V. Golovanov, V. V. Serdyuk, L. E. Stys, *et al.*, Ukr. Fiz. Zh. **33**, 390 (1988).
3. V. D. Bondar', A. B. Lyskovich, I. M. Matviishyn, and S. B. Kharambura, Izv. Akad. Nauk SSSR, Neorg. Mater. **26**, 660 (1990).
4. I. M. Matviishyn, V. D. Bondar', and A. B. Lyskovich, Vestn. L'vovsk. Univ., Ser. Fiz., No. 23, 27 (1990).
5. V. D. Bondar', A. B. Lyskovich, I. M. Matviishyn, and S. B. Kharambura, in *Proceedings of the 12th All-Union Conference on Physics of Semiconductors, Kiev, 1990*, Chap. 2, p. 112.
6. *Wide-Band Layered Crystals and Their Physical Properties*, Ed. by A. B. Lyskovich (Vishcha Shkola, Lvov, 1982) [in Russian].
7. *Fundamentals of Semiconductor Electronics*, Ed. by O. V. Snitko (Naukova Dumka, Kiev, 1985) [in Russian].
8. B. M. Kostyuk, A. B. Lyskovich, I. M. Matviishyn, and S. S. Novosad, Funct. Mater. **7**, 220 (2000).
9. M. A. Magomedov, Kh. A. Magomedov, M. A. Rizakhanov, and G. M. Gasanbekov, in *Proceedings of the 5th All-Union Conference on Physics and Technical Applications of Semiconductors, Vil'nyus, 1983*, Vol. 1, pp. 81–82.
10. V. G. Baru and F. F. Vol'kenshteĭn, *Effect of Radiation on Surface Properties of Semiconductors* (Nauka, Moscow, 1978) [in Russian].
11. F. F. Vol'kenshteĭn, *Electronic Processes on Semiconductor Surface at Chemisorption* (Nauka, Moscow, 1987) [in Russian].
12. V. A. Sokolov and A. N. Gorban', *Luminescence and Adsorption* (Nauka, Moscow, 1969) [in Russian].
13. G. P. Peka, *Physical Phenomena on Semiconductor Surface* (Vishcha Shkola, Kiev, 1984) [in Russian].
14. A. P. Akhoyan, N. E. Korsunskaya, and I. V. Markevich, Ukr. Fiz. Zh. **33**, 827 (1988).

Translated by A. Sidorova

Technical Physics, Vol. 50, No. 2, 2005, pp. 262–263. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 124–125.
Original Russian Text Copyright © 2005 by Balonishnikov.

BRIEF COMMUNICATIONS

Small-Scale-Velocity-Induced Linear Instability of Large–Scale Turbulent Flows

A. M. Balonishnikov
St. Petersburg State University of Economics and Engineering, St. Petersburg, 191002 Russia
e-mail: balonalex@yahoo.co.uk

Received January 8, 2004

Abstract—Analysis of a simplified equation derived previously for small-scale velocity components shows that any turbulent flow of an incompressible liquid becomes unstable against infinitesimal perturbations of small-scale velocity components if the strain rate tensor for the large-scale velocity is high. Such a statement comes into conflict with the classical stability theory, which specifically asserts that the Poiseuille flow in a circular tube is linearly stable against infinitesimal perturbations. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

The behavior of a turbulent flow under shear is of fundamental interest, since just this type of flow is most frequently encountered in applications. The question arises as to whether turbulence can be suppressed by the shear induced by a large-scale flow [1]. Our analysis based on an approximate equation for small-scale velocity components [2] shows that shear always enhances the turbulence in a linear approximation. On the other hand, data following from the analysis of the Orr–Sommerfeld equations indicate that the Poiseuille flow in a circular tube is linearly stable against infinitesimal velocity perturbations, however large the Reynolds number, being unstable against finite-amplitude perturbations [3, 4]. A possible physical explanation for such a contradiction will be given in the last section.

LINEAR ANALYSIS OF SMALL-SCALE POLARIZATION FOURIER COMPONENTS OF THE VELOCITY IN ANISOTRPIC TURBULENCE

With viscous forces showing up on small scales ignored, the instability depends on the positive real parts of characteristic indices $\lambda_{1,2}$ involved in the linear stability theory:

$$\lambda_{1,2} = -\frac{P}{2} \pm \frac{1}{2}\sqrt{P^2 - 4Q}, \tag{1}$$

where P contains only strain rate components $S_{ij} = \frac{1}{2}(\partial_j U_i + \partial_i U_j)$ (here, U_i is an ith component of the large-scale velocity) and Q contains, in addition, the components of large-scale vorticity vector $\mathbf{\Omega} = \mathrm{curl}\,\mathbf{U}$ [2].

Evidently, the condition $-P/2 > 0$ suffices to induce instability along some of the directions in the spectral space. Hereafter, we will seek for the directions along which quantity $-P/2$ is maximal. Note that the square root in expression (1) may only increase the value of λ_1. Rotating the coordinate system, one can bring symmetric tensor S to a diagonal form, with eigenvalues S_{11}, S_{22}, and S_{33} occupying the leading diagonal. They may be ranked in descending order as

$$S_{11} \ge S_{22} \ge S_{33}. \tag{2}$$

In view of the incompressibility condition, we have

$$S_{11} + S_{22} + S_{33} = 0. \tag{3}$$

In the spherical coordinate system, the expressions for P and Q after straightforward transformations take the form [2]

$$\begin{aligned} P = {} & \left(\frac{1}{4} - \frac{1}{4}\cos 2\theta - \frac{3}{4}\cos 2\eta - \frac{1}{4}\cos 2\eta\cos 2\theta\right)S_{11} \\ & + \left(\frac{1}{4} + \frac{1}{4}\cos 2\theta - \frac{3}{4}\cos 2\eta + \frac{1}{4}\cos 2\eta\cos 2\theta\right)S_{22}, \end{aligned} \tag{4}$$

$$\begin{aligned} Q = {} & -\frac{1}{4}(1 - \cos 2\theta)(1 + \cos 2\eta)S_{11}^2 \\ & - S_{11}S_2\cos 2\eta - \frac{S_{22}^2}{2}(1 + \cos 2\theta)(1 + \cos 2\eta) \\ & + \frac{1}{4}(\cos\theta\cos\eta\,\Omega_1 + \sin\theta\cos\eta\,\Omega_2 + \sin\eta\,\Omega_3)^2, \end{aligned} \tag{5}$$

where $\theta \in [0, 2\pi]$ and $\eta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ are the angles in the spherical coordinate system and $\mathbf{\Omega} = \mathrm{curl}\,\mathbf{U}$ ($\mathbf{\Omega} = \{\Omega_1, \Omega_2, \Omega_3\}$).

Let us introduce, for convenience, function

$$f = -\frac{P}{S_{11}}, \quad S_{11} > 0, \tag{6}$$

and seek its maximum in angular variables. In view of (2), the explicit expression for function f will be

$$\begin{aligned} f &= -\frac{1}{4}+\frac{1}{4}\cos 2\theta+\frac{3}{4}\cos 2\eta+\frac{1}{4}\cos 2\eta\cos 2\theta \\ &+\alpha\left(-\frac{1}{4}-\frac{1}{4}\cos 2\theta+\frac{3}{4}\cos 2\eta-\frac{1}{4}\cos 2\eta\cos 2\theta\right), \end{aligned} \tag{7}$$

where $\alpha = S_{22}/S_{11}$.

From (2) and (3), it follows that

$$\alpha \in \left[-\frac{1}{2}, 1\right]. \tag{8}$$

Now we introduce new variables $a = \cos 2\theta$ and $b = \cos 2\eta$ ($|a| \le 1$, $|b| \le 1$). The relationships

$$\partial_a f = \frac{1}{4}+\frac{1}{4}b-\frac{1}{4}\alpha-\frac{1}{4}b\alpha = 0, \tag{9}$$

$$\partial_b f = \frac{3}{4}+\frac{1}{4}a+\frac{3}{4}\alpha-\frac{1}{4}a\alpha = 0 \tag{10}$$

are the necessary maximum conditions for function f in the new variables.

At $\alpha = 1$, the linear set of Eqs. (9) and (10) has no solutions; at $\alpha \ne 1$, it has the unique solution

$$a = 3+\frac{6}{\alpha-1}, \tag{11}$$

$$b = -1. \tag{12}$$

From (8) and (11), it follows that

$$a \le -1. \tag{13}$$

With regard to the above restrictions imposed on variables a and b (i.e., to their ranges of definition), we conclude that, in the angular variables, function f may reach a maximum only at the extremities of their ranges of definition, i.e., on the sides of a unity square. Consider four sides of this square in detail:

(a) $a = 1, \quad -1 \le b \le 1,$

$$f = -\frac{\alpha}{2}+b\left(1+\frac{\alpha}{2}\right),$$

then, $f_{\max} = 1$ at $b = 1$;

(b) $a = -1, \quad -1 \le b \le 1,$

$$f = -\frac{1}{2}+b\left(\frac{1}{2}+\alpha\right),$$

then, $f_{\max} = \alpha$ at $b = 1$;

(c) $b = 1, \quad -1 \le a \le 1,$

$$f = \frac{1}{2}+\frac{\alpha}{2}+\frac{a}{2}(1-\alpha),$$

then, $f_{\max} = 1$ at $a = 1$;

(d) $b = -1, \quad -1 \le a \le 1,$

$$f = -1-\alpha,$$

then, f = const.

Thus, for any $\alpha \in \left[-\frac{1}{2}, 1\right]$, there exists a positive maximum of function f, $f_{\max} = 1$. In other words, in a coordinate system moving with local velocity U, any flow of an incompressible liquid that has a locally large-scale tensor $\mathbf{S} \ne 0$ becomes unstable against infinitesimal small-scale velocity perturbations provided that S_{11} is large enough for viscosity forces to be overcome [2]; that is,

$$\frac{S_{11}}{2} > \nu k_{\max}^2. \tag{14}$$

Here, ν is the molecular viscosity and $k_{\max} = 2\pi\sqrt{3}/L$, where L is a scale that separates large-scale and small-scale motions.

ANALYTICAL RESULTS ND CONCLUSIONS

Relationship (14) implies that, in a local coordinate system moving with large-scale velocity **U**, any large-scale flow of an incompressible liquid that exhibits arbitrary large-scale vorticity **Ω** is linearly unstable against infinitesimal perturbations of the small-scale velocity, provided that strain rate tensor **S** for the large-scale velocity is high (i.e., exceeds viscosity forces). Then, the validity of the approach adopted in the classical stability theory becomes questionable. This approach leads to the Orr–Sommerfeld equation, in terms of which, specifically, the parabolic profile of the Poiseuille flow in a circular tube reproduces the large-scale and small-scale regions of the flow simultaneously (the parabolic or linear profile contributes to the Fourier harmonics of small-scale vortices). Accordingly, the small-scale velocity components are taken into account twice. This raises the question: is the statement of the classical theory that the Poiseuille laminar flow is stable against infinitesimal velocity perturbations [3, 4] correct? It would be of interest to experimentally verify the results obtained in this paper, relationships (14) in the first place, by artificially generating small-scale harmonic perturbations like those considered in [4].

REFERENCES

1. P. W. Terry, Rev. Mod. Phys. **72**, 109 (2000).
2. A. M. Balonishnikov, Zh. Tekh. Fiz. **72** (10), 106 (2003) [Tech. Phys. **48**, 1255 (2003)].
3. P. G. Drazin and W. H. Reid, *An Introduction to Hydrodynamic Stability Theory* (Cambridge Univ. Press, Cambridge, 1980).
4. B. Hof, A. Juel, and T. Mullin, Phys. Rev. Lett. **91**, 244502 (2003).

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 264–266. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 126–128.
Original Russian Text Copyright © 2005 by Minaev, Rukhadze.

BRIEF COMMUNICATIONS

Effect of a Plasma Layer on the Vortex Structures in a Gas Flow

I. M. Minaev and A. A. Rukhadze
Prokhorov Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 119991 Russia

Received February 5, 2004

Abstract—Questions related to the effect of a plasma layer on the vortex structures in a gas flow are considered for the case in which the layer forms in a turbulent flow or in a flow of a mixed type. It is shown that the onset of a plasma layer in the turbulent flow leads to the efficient suppression of the low-frequency components of the turbulent spectrum and to the displacement of the spectrum toward higher frequencies. *© 2005 Pleiades Publishing, Inc.*

When a viscous gas flows along the surface of a thin plate, a laminar or a turbulent boundary layer (or a layer of mixed type) arises near the plate. In the case of a turbulent flow or a flow of a mixed type, the flow structure at the outer boundary of a viscous sublayer in an inhomogeneous turbulent region near the plate surface is spontaneously destroyed. The substantial vorticity generated at the plate surface evolves into pronounced vortices, which then move away from the surface sublayer and carry with them the angular momentum into the outer region [1]. In statistical models [2] used to describe the physical processes occurring in developed turbulence, it is assumed that such a flow is generated by an ensemble of irregularly evolving vortex elements (perturbations, inhomogeneities, etc.) with very different sizes. In order of magnitude, the sizes of the largest vortices are equal to the dimension of the region of turbulent motion and the sizes of the smallest vortices are equal to the dimension of the region across which the momentum can be efficiently transported under the action of molecular viscosity, which smoothes the velocity gradients [3].

Some questions related to a viscous gas stream flowing around a plasma sheet produced by a slipping discharge were discussed in [4, 5]. In [4], the problem was analyzed of how the plasma sheet affects the character of the gas flow, as well as of the dependence of the viscosity on the degree of gas ionization in the plasma sheet. In [5], the issues regarding the dynamic evolution of a slipping discharge at the surface of a dielectric plate in a viscous gas flow were considered. The results obtained in those papers show that, after the discharge has come to an end, a plasma layer with a characteristic temperature distribution over its cross section forms in a boundary layer above the plate around which the gas flows [5]. Figure 1 shows the profile of the normalized temperature $T/T_{\max}$ across the boundary layer at the time t_0 just after the termination of energy deposition in the slipping discharge (in the figure, δ^* corresponds to the maximum temperature $T_{\max}$ [5]).

In order to gain insight into the effect of the plasma region on the vortex structures, it is important to determine the relationship of the spatial scale of the smallest vortices to the Debye screening length and the molecular mean free path because it might be necessary to take into account molecular motions in order to describe situations in which the motions of the medium occur on very short spatial scales. In turbulent flows at $P_0 \approx 1$ atm, the average molecular mean free path is much less than the sizes of the vortices [2, 3] and, in the case under consideration, it is also less than the Debye screening length ($n \approx 10^{12}$ cm^{-3}, $R_D \approx 2.5 \times 10^{-4}$ cm), so we can apply a fluid model [2]. If we consider a vortex as a tube (Fig. 2), associate the angular velocity ω_V with the frequency f_V, and define the period of the vortex by the equality $T_V = 2\pi/\omega_V$, then we can speak of a turbulent flow spectrum whose width may amount to 10 kHz and even more [3]. It is usually assumed that vortices of vastly different sizes only slightly influence one

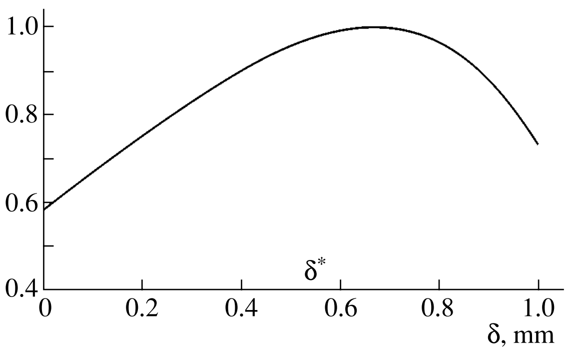


Fig. 1. Profile of the normalized temperature $T/T_{\max}$ over the cross section of the boundary layer for $T_{\max} \approx 3000$ K, $\delta^* \approx 0.7$ mm, and $\delta = 1$ mm.

another; it is only vortices of comparable size that can exchange energy. It was established that viscosity has an insignificant impact on the motion and structure of the main turbulent flow, but it plays a governing role in the final stage of turbulent energy dissipation, when the velocity gradients in small vortices are smoothed out by viscous stresses [2]. If a plasma region arises within the turbulent flow, then the energy of large-scale vortices (such that their sizes are comparable to the transverse size δ of the plasma region, see Fig. 2) can be lost mainly because of the rotation of an electric dipole formed by a vortex in the nonuniform electric field of the plasma layer.

Let us examine this process in more detail. We assume that, after the slipping discharge has come to an end, it leaves a plasma region with a charged particle density (at the time t_0) of about $n \approx 10^{12}$ cm^{-3} [6]. At later times, the plasma density decreases as a result of diffusion toward the boundaries of the region. After the plasma has relaxed to a quasineutral state (the characteristic relaxation time being $\tau_M \sim (R_D)^2/4D_e$, where D_e is the electron diffusion coefficient [7]), the plasma region of thickness δ is dominated by the ambipolar diffusion of charged particles toward its outer boundaries. Let there be a vortex with a characteristic angular velocity ω_V and characteristic radius R_V ($R_V < \delta^*$, see Fig. 2) within a plasma layer of thickness δ^*. We assume that the density distribution of charged particles over a cylinder of unit length with radius R_V corresponds to the density distribution of charged particles inside the layer. In the one-dimensional case, the diffusive flow of charged particles is directed along the normal to the wall (which is assumed to be in the XZ plane). The cylinder rotates in the direction indicated in Fig. 2. In this case, under the assumption that the cylinder rotates as a single entity, the lower density regions lying inside the rotating cylinder will fall into higher density regions lying outside the cylinder. In this case, the charged particles will diffuse from the outer regions (where their density is higher) into the regions inside the cylinder (where their density is lower). The depth to which the charged particles diffuse into the cylinder is determined by the rotation velocity of the cylinder, the electron diffusion velocity, the ambipolar diffusion velocity, and the gradients of the charged particle densities between the regions outside and inside the cylinder.

The accumulation of the electric charge at the surface of the rotating cylinder results from diffusion caused by the difference in the charged particle densities in the outer and inner layers just adjacent to the surface of the cylinder. The depth to which the charged paticles diffuse into the rotating vortex does not exceed the Debye radius R_D. Let us estimate the charge accumulated at the cylinder surface. This charge is determined by the ambipolar electric field E_A and also by the ratio of the linear velocity u_l of the surface of the rotating cylinder to the diffusion velocity u_A of the charged particles. The temperature distribution over the cross section of the boundary layer was calculated in our earlier paper [5]. According to the results of those calculations, we have $\nabla n/n \gg \nabla T/T$, which indicates that thermal diffusion is unimportant. The ambipolar electric field E_A is described by the expression [8]

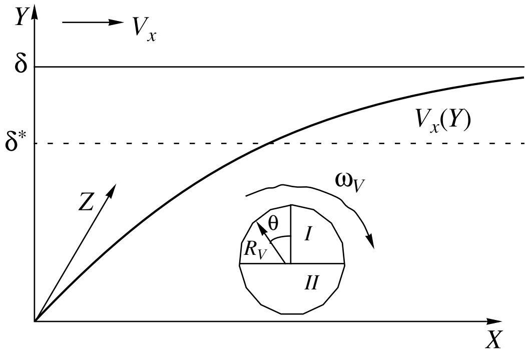


Fig. 2. Schematic cross section illustrating the position of the vortex in a plasma layer of thickness δ.

$$E_A \approx (T/e)\nabla n/n, \quad (1)$$

and the ambipolar diffusion velocity u_A is equal to

$$u_A = -D_A \nabla n/n, \quad (2)$$

where $D_A = 2T/m_i\nu_{in}$ is the ambipolar diffusion coefficient, m_i is the mass of an ion, ν_{in} is the ion–neutral collision frequency (the plasma is assumed to be weakly ionized and isothermal, $T_e = T_i = T \approx 0.3$ eV), and e is the charge of an electron.

Under these circumstances, the electron–neutral collision frequency in air is equal to $\nu_{en} = 4 \times 10^9 P_0$ s^{-1} and the ion–neutral collision frequency is equal to $\nu_{in} = (\sqrt{m/M})\nu_{en} = 2 \times 10^7 P_0$ s^{-1}, where P_0 is the air pressure in torr [9].

Using formulas (1) and (2), we can determine the dipole moment of a vortex in an inhomogeneous plasma layer. The extent of charge separation in the ambipolar field E_A in region I in Fig. 2 is determined by the depth l to which the charged particles diffuse into the inner layer just adjacent to the surface of the rotating vortex. For $\omega_V = 0$, the depth l is on the order of $R_D(E_A/E_D)$, where E_D = T/eR_D. Under the conditions adopted above, we have $R_D \approx 2.5 \times 10^{-4}$ cm. For $n \approx 10^{12}$ cm^{-3}, this yields the following value of the maximum surface charge density: $\sigma_0 \approx 6.3 \times 10^5 e$ cm^{-2}. The charge density profile in the plasma layer is given by the expressions [8]

$$n = n_{\max}\cos(X\sqrt{\nu_i/D_A}), \quad \delta^*\sqrt{\nu_i/D_A} = \pi/2, \quad (3)$$

where $\nu_i \approx \pi^2 D_A/4(\delta^*)^2$ is the rate of ionizing collisions [8]. The angular distribution of the surface charge den-

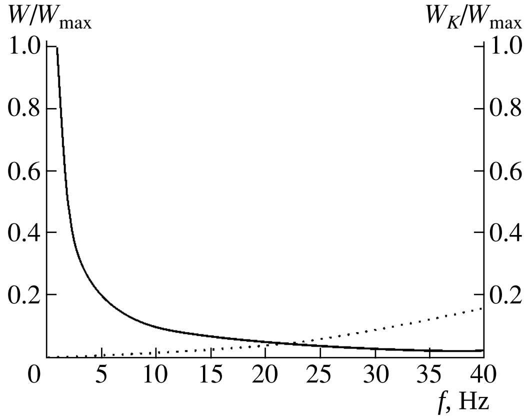


Fig. 3. Dependence of the quantities W (solid curve) and W_K (dashed curve) on f for a fixed value of R_V.

sity over the surface of the cylinder has the form $\sigma = \sigma_0 \cos\theta$.

In the ambipolar electric field E_A, a vortex having such a surface charge density distribution rotates with the angular velocity ω_V. As a result of this rotation, the dipole moment of the vortex is averaged out. Under our conditions, we have $E_A \approx 2.6 \times 10^{-2}$ V/cm, $\nu_{en} \approx 3 \times 10^{12}$ s^{-1}, $\nu_{in} \approx 1.5 \times 10^{10}$ s^{-1}, $D_A \approx 2D_i \approx 7.2 \times 10^{-1}$ cm^2/s, $D_e \approx 78$ cm^2/s, $\sigma_0(\max) \approx 6.9 \times 10^5 e$ cm^{-2}, and $\nu_i = 7 \times 10^2$ s^{-1}.

The dipole moment P is equal to [10]

$$P = E_A \sigma_0 R_V^2 / 2. \qquad (4)$$

A dipole having such a moment rotates in the ambipolar electric field E_A of the plasma layer; in this case, the dipole moment is inversely proportional to the rotation velocity ω.

Hence, a steady distribution of the charges forming the dipole is established at the expense of the energy of the rotating vortex. Let us estimate the energy lost by a vortex with a dipole moment P during its rotation in the ambipolar electric field E_A. Provided that the viscous friction can be ignored, the portion W of the kinetic energy lost by the rotating vortex is determined by the energy lost by the dipole rotating in the field E_A. The energy lost during the time that a vortex of radius $R_V = \delta/2$, rotating with an angular velocity $\omega = 1$ rad/s, executes one revolution is equal to $W_{\max} \approx 10^{-14}$ J. The kinetic energy of a vortex of the same size is equal to $W_K \approx 2.5 \times 10^{-18}$ J. The dependence of the energies W_K and W on $f = \omega/2\pi$ is illustrated in Fig. 3.

The results obtained above show that (i) the kinetic energy lost by a vortex in a plasma layer is determined by the dipole moment and angular velocity, (ii) the energy lost by a vortex of fixed radius decreases in inverse proportion to the angular velocity, and (iii) the kinetic energy of a vortex increases in direct proportion to the square of the angular velocity. Hence, the formation of a plasma layer in the turbulent flow region leads to the efficient suppression of the low-frequency components of the turbulent spectrum, thus resulting in the displacement of the spectrum toward higher frequencies, i.e., toward the frequency range in which the energy of the vortices is efficiently dissipated by viscous friction.

ACKNOWLEDGMENTS

We are grateful to L.D. Tsendin for critical comments and for much useful advice.

REFERENCES

1. N. F. Krasnov, *Aerodynamics* (Vysshaya Shkola, Moscow, 1980) [in Russian].
2. W. Frost and T. H. Moulden, *Handbook of Turbulence* (Plenum, New York, 1977; Mir, Moscow, 1980).
3. L. D. Landau and E. M. Lifshitz, *Mechanics of Continuous Media* (Nauka, Moscow, 1982).
4. G. P. Kuz'min, I. M. Minaev, and A. A. Rukhadze, Teplofiz. Vys. Temp. **40**, 515 (2002).
5. I. M. Minaev and A. A. Rukhadze, Teplofiz. Vys. Temp. **41**, 806 (2003) [High Temp. **41**, 731 (2003)].
6. I. O. Kovalev, G. P. Kuz'min, and A. A. Nesterenko, Tr. IOFAN **52**, 52 (1996).
7. V. A. Rozhanskiĭ and L. D. Tsendin, *Collision Transfer in Partially Ionized Plasma* (Énergoatomizdat, Moscow, 1988) [in Russian].
8. V. N. Golant, A. P. Zhilinskii, and I. E. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).
9. S. Brown, *Basic Data of Plasma Physics: The Fundamental Data on Electrical Discharges in Gas* (AIP, New York, 1993; Gosatomizdat, Moscow, 1980).
10. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).

Translated by O. Khadin

Technical Physics, Vol. 50, No. 2, 2005, pp. 267–269. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 129–130.
Original Russian Text Copyright © 2005 by Kalinin, Solomein, Aleksandrov, Atanov.

BRIEF COMMUNICATIONS

Effect of the Isotope Composition and Unsaturated Bonds on the Mass Defect Spectrum of Residual Hydrocarbons

B. A. Kalinin, A. A. Solomein, O. E. Aleksandrov, and V. E. Atanov

Ural State Technical University (UPI), ul. Mira 19, Yekaterinburg, 620002 Russia

e-mail: aleks@dpt.ustu.ru

Received April 20, 2004

Abstract—By accurately measuring the position of the peaks in the ion mass spectrum of residual gases, it is found that the peaks are associated with hydrocarbon ions. The “mass defect” spectrum is close to that obtained analytically. The deviations observed are related to the presence of the ^{13}C isotope and nonsaturation of bonds in the ions recorded. The experimental data obtained may be used as a reference in impurity microanalysis, in calibrating the mass scale of mass spectrometers, and in checking the theory of mass spectrum analysis. © *2005 Pleiades Publishing, Inc.*

The sensitivity of commercial spectrometers produced in the Commonwealth of Independent States allows direct (without concentration or using reference mixtures) analysis of microimpurities at a level of 1–10 ppm. Impurity ion identification is facilitated in the presence of reference masses (peaks), as which some of the residual gas peaks may be used. The positions of the remaining peaks must also be known, since they may merge with the impurity peaks.

The basic goal of this work is to identify the mass spectra of residual gases in the mass interval 300–500, which partially covers the mass spectrum of fragments of a uranium hexafluoride molecule or a molecule of heavy partially or completely substituted haloid hydrocarbons. Calibration of the mass scale in this case is similar to the calibration used in determining the masses of metastable uranium hexafluoride ions [1].

The measurements were made with an MI-1201 AGM mass spectrometer in the ion count regime, which was equipped with a secondary emission multiplier and did not need any hardware refinement. Since the instrument is delivered by the vendor with a control and data acquisiton program compatible with MS-DOS, we devised a program with similar functionality in the software environment Delphi for the operational system Windows 9x/NT/2000. In particular, the program allows for determination of the position of a peak and its width at a given height and mass calibration using two peaks. The standard service of the software environment allows one to look through spectrum fragments. Among the drivers, there is the one starting the program in MS-DOS. The Windows operational systems make network control of the instrument and program authorization (in the case of Windows NT/2000) possible.

The mass spectrum may be stored as a file written in the format of the program or as a text file containing two columns of figures.

The processing of the mass spectrum, including the determination of the position, heights, and widths of the peaks; the background level; and the associated errors, is accomplished by another (processing) program [2]. The processing program splits the mass spectrum into peak-containing areas and approximates each of the peaks by a Gaussian function with the least squares method. It also allows correction of the mass scale using known (reference) lines. Also, the time of processing of the mass spectrum is calculated.

Both programs complemented by the user’s manual and the mass spectrum considered in this work are available from the site http: //www.mp.dpt.ustu.ru./mass spectrometry/.

The measurements were performed without cooling the traps of the magnetic discharge pumps at a pressure of ≈3×10^{-6} Pa (the reading of a vacuum gage), an emission current from the cathode of 0.1 mA, and an ionization voltage of 50 V. For the ion masses to be measured with a high accuracy, the mass spectrum was recorded with a step of 0.01 u and an integration time of 0.1 s under a continuously varying magnetic field. The basic reason for the systematic drift of the mass scale was probably heating of the uncooled magnet during its long-term operation in the large mass range.

1063-7842/05/5002-0267$26.00 © 2005 Pleiades Publishing, Inc.

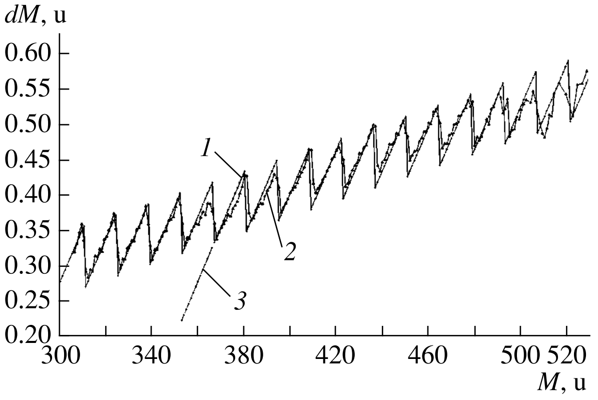


Fig. 1. Measured and calculated defect mass spectra for residual hydrocarbons. *M*, mass; *dM*, mass defect; *1*, calculated spectrum; *2*, measured spectrum; and *3*, ions with a high concentration of unsaturated bonds.

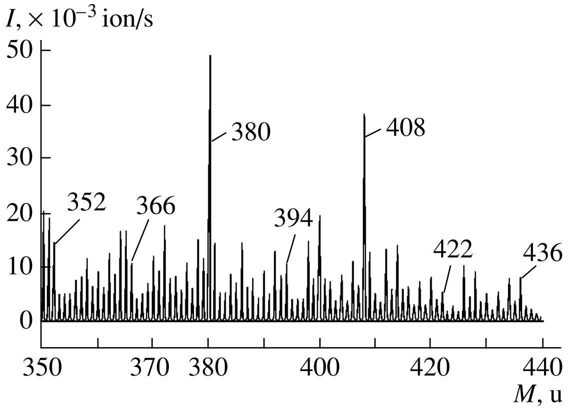


Fig. 2. Fragment of the mass spectrum of residual hydrocarbons. The positions of the molecular ions are indicated.

To identify the mass spectrum, specifically to locate the molecular ion peaks, it was processed by the program mentioned above [2]. The mass scale was calibrated against the highest and narrowest peaks, which were 3 u distant from the molecular peaks on either side (the reason for such an effect will be clarified below).

The calibration (correction) curve for the positions of the peaks that were calculated by the program was approximated by a parabola with a standard deviation not exceeding 10 mmu (1 mmu = 10^{-3} u). Then, the mass spectrum was refined (corrected) by taking into account the parameters of the parabola.

The position of the peaks in the mass spectrum thus obtained is convenient to describe by introducing the notion of ion mass defect (see [3, 4, p. 76]). The mass defect is defined as the difference between the mass expressed in atomic mass units and the mass number. The mass defect spectra of residual hydrocarbons (one was derived (measured) from the mass spectrum corrected and the other was calculated using the tabulated values of the masses, which ignore the isotope composition) are presented in Fig. 1. The sharp variation of the mass defect in the spectrum measured is due to changes in the ion composition,

$$C_xH_y \longrightarrow C_{x+1}H_{y-11}. \tag{1}$$

Figure 1 also shows a fragment of the spectrum calculated for $C_{27}H_x$ ions where the concentration of unsaturated bonds exceeds the minimal necessary value for the given mass.

The qualitative agreement of the spectra indicates the adequate identification of the mass spectrum. The increasing and more irregular discrepancy observed at masses above 480 u is due to the depression of the peaks. The discrepancy is most pronounced near the molecular ion masses. This is because most of the molecular ions constituting the peaks next to the molecular peaks (on the side of larger masses) contain the ^{13}C isotope; therefore, the points of the mass defect spectrum that correspond to these peaks lie above the spectrum calculated. In this case, change (1) in the isotope composition turns into

$$^{12}\mathrm{C} \longrightarrow {}^{13}\mathrm{C}. \tag{2}$$

For example, peaks 381 and 409 (Fig. 2) almost completely consist of the molecular ions $^{13}C^{12}C_{26}H_{56}$ and $^{13}C^{12}C_{28}H_{60}$, respectively. For these ions, the deviation from the spectrum calculated is 85 mmu. The height of ^{13}C isotope-containing peak 381 that was calculated from the height of peak 380 (with allowance for the isotopic contribution of the latter to peak 379) coincides with the measured height (accurate to the measurement error). This means that peak 381 consists only of molecules containing the ^{13}C isotope. That is why the mass defects for peaks 380 and 381 are close to each other.

The fact that the molecular peaks with masses 380 and 408 are high (which indicates the stability of $C_{27}H_{56}$ and $C_{29}H_{60}$ molecules against electron impact) needs special consideration. The interplay between the peaks discussed above shows up not only near the molecular ions but also when the change $^{12}C \longrightarrow {}^{12}CH$ in the ion composition (the mass defect increases by 7.8 mmu) gives way to the change $^{12}C \longrightarrow {}^{13}C$ (the increase by 4.5 mmu).

Ions with a high concentration of unsaturated bonds (Fig. 1) basically may greatly influence the mass defect. In our case, however, such ions are present in small amounts, as follows from the moderate deviations from the calculation data, including the deviation related to the change in the isotope composition. The influence of such ions is yet seen in the reduction of the mass defects for the peaks lying on the other (smaller mass) side of the molecular peaks. The reduction is distinct, e.g., near peak 366. One more piece of evidence that the ions with a high concentration of unsaturated bonds are contributors to the mass defect is the broadening of the peaks subject to influence of such ions.

Therefore, when calibrating the mass scale of a mass spectrometer, one should select high narrow peaks far away from molecular peaks. If, in our case, the calibration masses met the formula M_nH_{2n-6}, the standard deviation of the calibration (correction) curve would have reduced twofold to become 5 mmu.

REFERENCES

1. B. A. Kalinin, V. E. Atanov, and O. E. Aleksandrov, Zh. Tekh. Fiz. **72** (5), 135 (2002) [Tech. Phys. **47**, 648 (2002)].
2. A. A. Solomein, B. A. Kalinin, and L. M. Aleksandrov, Analitika Kontrol’ **7**, 35 (2003).
3. I. G. Zenkevich and B. V. Ioffe, *Identification of Mass Spectra of Inorganic Compounds* (Khimiya, Leningrad, 1986) [in Russian].
4. *Advances in Mass Spectrometry*, Ed. by J. D. Waldron (McMillan, New York, 1959; Inostrannaya Literatura, Moscow, 1963).

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 270–273. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 131–134.
Original Russian Text Copyright © 2005 by Rybka, Baksht, Lomaev, Tarasenko, Krishnan, Thompson.

BRIEF COMMUNICATIONS

Emission Characteristics of a Pulsed Discharge in Xenon

D. V. Rybka*, E. Kh. Baksht*, M. I. Lomaev*, V. F. Tarasenko*,
M. Krishnan, and J. Thompson****

** Institute of High Current Electronics, SB RAS, Tomsk, 634055 Russia*
*** Alameda Applied Sciences Corporation, CA 94577, San Leandro, USA*
e-mail: VFT@loi.hcei.tsc.ru

Received May 25, 2004

Abstract—Emission characteristics of a high-current pulsed discharge in xenon are studied experimentally. The study is aimed at developing a source of spontaneous UV radiation (with $\lambda \leq 250$ nm) for controlling high-voltage crystalline diamond switches. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

Spontaneous emission sources based on pulsed or continuous-wave discharges in gases or gas–vapor mixtures have found widespread application [1–4]. Of great promise are pulsed UV lamps based on a freely expanding high-pressure discharge (the so-called globular pulsed lamps) [4]. Distinctive features of such lamps are the short pulse duration, the high emission power, a fairly high (a few electronvolts) temperature of the discharge plasma, the broad emission spectrum containing a continuum component, and the low discharge volume that enables the efficient optical focusing of radiation onto an irradiated object. As compared to discharges in other noble gases, discharges in xenon are characterized by the highest electric field and the lowest potential drop across the electrode sheaths. This feature makes discharges in xenon most promising from the standpoint of the lamp efficiency [1]. It is reasonable to employ the advantages of xenon globular lamps in developing inexpensive efficient emission sources for controlling high-voltage crystalline diamond switches [5–7]. It has been shown that such switches can be controlled by an electron beam [8] or a UV laser [9]. However, the use of lasers or electron accelerators in commercial switches is inexpedient because of their high cost. In this context, the development of an inexpensive pulsed UV source for controlling a crystalline diamond switch is a challenging problem. It is necessary that the bulk of the emission energy of such a source lie in the wavelength range of $\lambda \leq 250$ nm, which corresponds to the fundamental absorption band of crystalline diamond ($\lambda \leq 225$ nm), as well as to the impurity absorption band [10, 11]. It is also desirable that the pulse duration of such a source be no longer than a few microseconds.

This paper, which is a continuation of [12], is devoted to studying the emission characteristics of the plasma of a freely expanding discharge in xenon.

EXPERIMENTAL SETUP AND MEASUREMENT TECHNIQUE

The experimental setup consisted of a storage capacitor bank and a high-voltage generator loaded on a pulsed gas-discharge lamp. The cylindrical quartz lamp with an inner diameter of 20 mm was filled with xenon. The interelectrode distance was 5 mm. The transmittance of the lamp wall in the spectral range of 200–250 nm was no less than 85%.

The discharge circuit consisted of pulsed lamp *F*, storage capacitor C_0, and trigatron switch *S* (Fig. 1). Capacitor C_0 was charged through resistance *R*. The charging voltage U_0 was 12 kV; the storage capacitance C_0 was either 3.3 or 233 nF; and the natural oscillation period of the discharge circuit was 0.08 or 0.85 μs, respectively.

The system for recording the discharge emission consisted of an FÉK-22SPU coaxial phototube and an EPP2000C-25 spectrometer (manufactured by StellarNet Inc., USA) equipped with a CCD array photodetector. The spectrometer was used to record the emission spectrum (in relative units) in the 200- to 850-nm spectral range. The FÉK-22SPU, with the known absolute spectral sensitivity in the 200–650 nm spectral range,

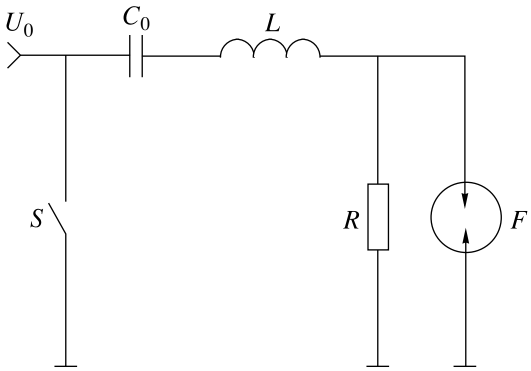


Fig. 1. Discharge circuit.

was used to monitor the time evolution of the emission power. The combined data from the EPP2000C-25 and FÉK-22SPU allowed us to record the emission spectrum in absolute units within the spectral sensitivity range of the FÉK-22SPU.

The discharge glow was photographed with a SensiCam CCD camera. To synchronize the discharge ignition with the operation of the CCD camera, the discharge gap was illuminated with radiation from a weak spark discharge in air. The discharge current and voltage were monitored with a current shunt and a resistive voltage divider, respectively.

A diamond switch was modeled by a IIa-type crystalline diamond photodetector manufactured by the Alameda Applied Sciences Corporation. The electric circuit of the photodetector was similar to that used in [9] and consisted of a detector, a storage capacitor, and a load resistance. The charging voltage of the storage capacitor was 250 V.

EXPERIMENTAL RESULTS AND DISCUSSION

In our experiments, we recorded the waveforms of the discharge current and voltage, the emission power and spectrum of the discharge, and the time evolution of the visible discharge glow.

Figure 2 shows the waveforms of the current and voltage for a discharge in xenon (for C_0 = 233 nF and U_0 = 12 kV) at a pressure of 550 Torr, as well as the time evolution of the emission power from the discharge and the power deposited in the discharge. It can be seen that more than 70% of the deposited energy is supplied in the first half-period of the discharge current oscillations. About 36% of the energy stored in the capacitor C_0 was deposited in the discharge, whereas the rest of energy was dissipated in the switch. The emission energy within the wavelength range of 200–250 nm was 4.2% of the energy deposited in the discharge. The total emission energy within the 200- to 850-nm wavelength range under study was ~1 J, the peak emission power being ~500 kW. The full width at half-maximum (FWHM) of the discharge emission pulse was 1.6 μs. The emission power reached its maximum ~700 ns after the discharge ignition. The discharge afterglow lasted over a few microseconds.

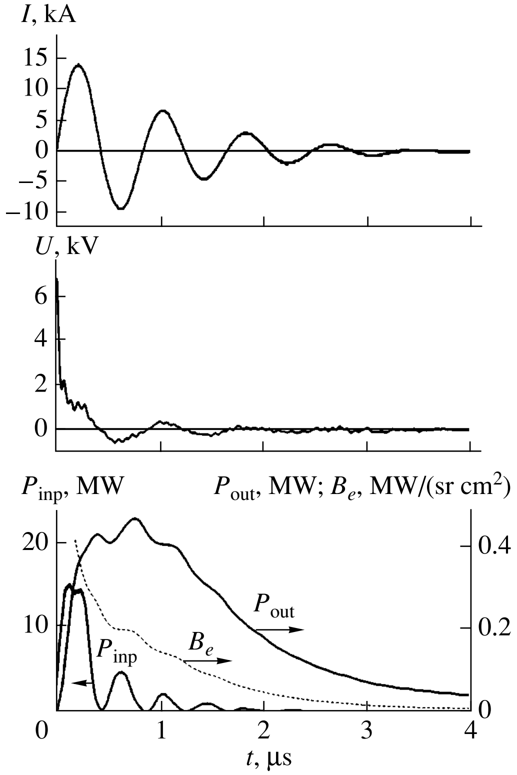


Fig. 2. Waveforms of the discharge current I, discharge voltage U, emitted power P_{out}, discharge brightness B_e, and deposited power P_{inp}.

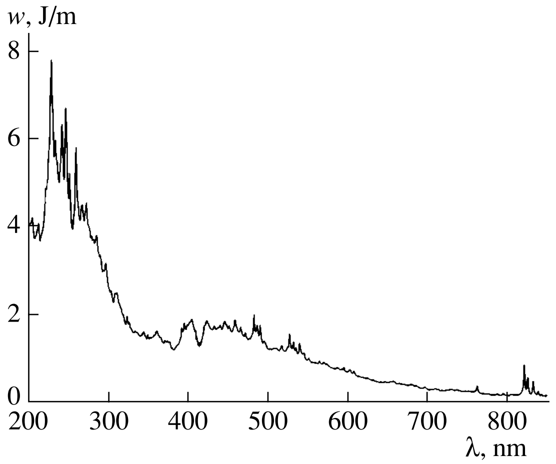


Fig. 3. Energy spectrum of the discharge emission.

The discharge emission spectrum is shown in Fig. 3. It can seen that most of the emission energy falls into the UV region.

Photographing a discharge in xenon (for $\tau \approx 0.85$ and 0.08 μs) with a CCD camera (see Fig. 4) showed that the breakdown of the discharge gap was almost always multichannel. Individual channels then expanded and, after ~200 ns, merged together.

As is known [3], the spatial distribution of radiation from pulsed lamps is mainly determined by the distribution of the emission intensity over the plasma volume. Therefore, it is of interest to study the time evolution of the plasma glow.

Figure 4a illustrates the time evolution of the discharge glow at C_0 = 233 nF and p = 550 Torr. It can be seen that the discharge glow becomes more uniform as the glowing region expands.

For a much smaller storage capacitance, the plasma glow remained spatially nonuniform till the end of the

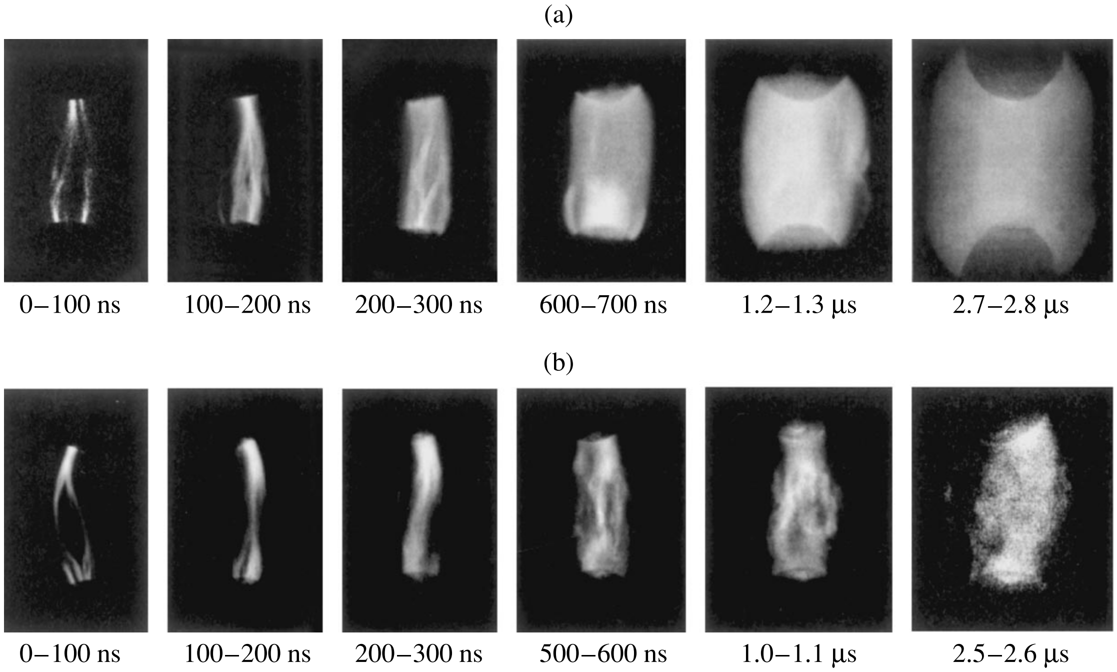


Fig. 4. Photographs of a discharge for C_0 = (a) 233 and (b) 3.3 nF.

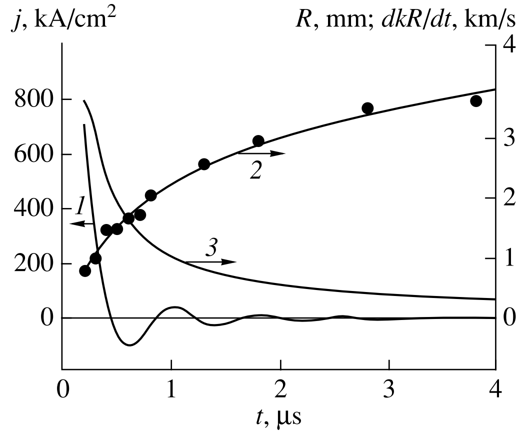


Fig. 5. Time evolution of (*1*) the discharge current density, (*2*) the discharge channel radius, and (*3*) the expansion rate of the discharge channel. The circles show the measured discharge channel radii.

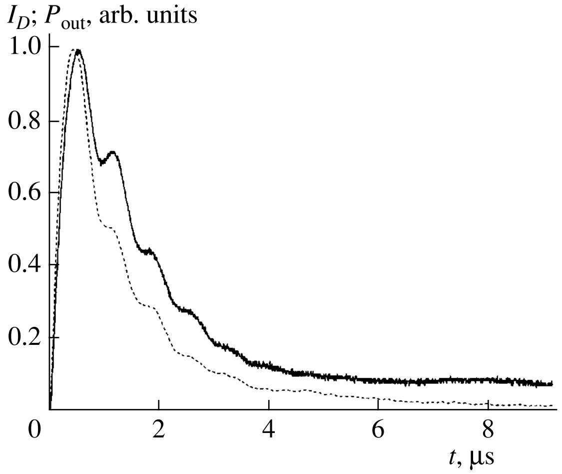


Fig. 6. Waveforms of the detector current I_D (dotted curve) and discharge emission power P_{out} (solid curve).

glow. This can be seen in Fig. 4b, which shows a series of photographs of a discharge for C_0 = 3.3 nF, U_0 = 12 kV, and p = 550 Torr. The FWHM of the glow is 220 ns. The glow intensity is maximal within the first 150 ns, when there are several discharge channels.

The photographs of the discharge glow allowed us to estimate the radius of the discharge channel, the rate at which it expands, and the averaged (over the discharge cross section) current density (Fig. 5), as well as to determine the time behavior of the discharge brightness. The estimates for C_0 = 233 nF are shown in Fig. 2. These results are valid starting from the instant $t \sim$ 200 ns, after which there is only one discharge channel (the discharge brightness and the current density were calculated using the approximate time dependence of the discharge channel radius).

When a crystalline diamond detector was illuminated with the discharge radiation (at C_0 = 233 nF, τ = 1.1 μs, and p = 550 Torr), the waveform of the current pulse in the detector circuit differed from the waveform of the emission pulse (Fig. 6). The reason is that the time evolution of the emission power in the spectral range of $\lambda \leq 300$ nm (recorded by the detector) was not the same as that in the wavelength range of 200–650 nm (recorded by the FÉK-22SPU). In [13], under similar conditions, the emission pulse in the wavelength range of 200–300 nm was also observed to be shorter than that in the range of 200–650 nm.

CONCLUSIONS

We have studied the spectral, energy, and temporal characteristics of a pulsed discharge in xenon. The time evolution of the discharge glow has been studied using a CCD camera. At a pressure of 550 Torr, the emission energy in the 200- to 850-nm wavelength range was ~1 J, and in the 200- to 250-nm wavelength range, it was ~0.3 J, the conversion efficiency of electric energy into radiation being 4.2%. The maximal discharge brightness was >400 kW/(sr cm^2), and the peak emission power was ~500 kW. The results obtained open up prospects of using pulsed xenon lamps to control high-voltage crystalline diamond switches.

ACKNOWLEDGMENTS

This study was supported by the US Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF) (grant no. RP2-538-TO-02) and NATO (grant no. EST.CLG.9495401).

REFERENCES

1. G. N. Rokhlin, *Discharge Sources of Light* (Énergoatomizdat, Moscow, 1991) [in Russian].
2. *Handbook on Illumination Engineering*, Ed. by Yu. B. Aĭzenberg (Énergoatomizdat, Moscow, 1983) [in Russian].
3. *Pulsed Light Sources*, Ed. by I. S. Marshak (Énergiya, Moscow, 1978) [in Russian].
4. *Encyclopedia of Low-Temperature Plasma*, Vol. 4: *Gas and Plasma Lasers*, Ed. by V. E. Fortov (Nauka, Moscow, 2000) [in Russian].
5. P. K. Bharadwaj, R. F. Code, H. M. van Driel, and E. Walentynowicz, Appl. Phys. Lett. **43**, 207 (1983).
6. P. T. Ho, C. H. Lee, J. C. Stephenson, and R. R. Cavanagh, Opt. Commun. **46**, 202 (1983).
7. J. Glinski, X.-J. Gu, R. F. Code, and H. M. van Driel, Appl. Phys. Lett. **45**, 260 (1984).
8. R. R. Prasad, S. W. Gensler, N. Qi, M. Krishnan, and G. Loubriel, in *Proceedings of the Conference on SAE Aerospace Power Systems, Mesa, Arizona, 1999*, p. 193.
9. E. I. Lipatov, A. N. Panchenko, V. F. Tarasenko, *et al.*, Kvantovaya Élektron. (Moscow) **31**, 1115 (2001).
10. J. E. Field, *The Properties of Diamond* (Academic, London, 1979).
11. G. B. Bokiĭ, G. N. Bezrukov, Yu. A. Klyuev, *et al.*, *Natural and Synthesized Diamonds* (Nauka, Moscow, 1986) [in Russian].
12. M. I. Lomaev, D. V. Rybka, V. F. Tarasenko, *et al.*, Izv. Vyssh. Uchebn. Zaved. Fiz., No. 1, 81 (2004).
13. A. A. Lisenko, M. I. Lomaev, D. V. Rybka, and V. F. Tarasenko, Proc. SPIE **4977**, 434 (2003).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 50, No. 2, 2005, pp. 274–275. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 135–136.
Original Russian Text Copyright © 2005 by Ganzherli, Denisyuk, Maurer, Chernykh.

BRIEF COMMUNICATIONS

On-Axis Diffuse Screen Based on a Reference-Free Volume Hologram

N. M. Ganzherli, Yu. N. Denisyuk, I. A. Maurer, and D. F. Chernykh

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia
e-mail: nina@holo.ioffe.rssi.ru

Received May 27, 2004

Abstract—A new type of hologram that elaborates upon the simplest (Gabor) on-axis hologram is suggested. The new approach makes it possible to eliminate the projected image distortions typical of the Gabor hologram. Specifically, based on a reference-free thick hologram, an on-axis holographic screen that does not transmit the zeroth order is prepared. In addition, this screen does not produce a halo and the conjugate image. It allows for recording in one spectral range and reconstruction in another, thereby greatly simplifying the choice of a light-sensitive record medium. With this screen, a color image can be projected. *© 2005 Pleiades Publishing, Inc.*

Elaborating upon the concept of recording the Gabor hologram for creating image-projecting screens, we have conceived the idea of using a reference-free volume hologram for this purpose. The essence of the idea is that, during recording the reference-free hologram of a diffuser, phase modulation may suppress the zeroth order in the image reconstructed. The scheme for recording such a screen is depicted in Fig. 1. Diffuser *D* is illuminated by coherent radiation $\bar{l}_0$. When interfering, the rays scattered from points P_1, P_2, P_3, … of diffuser *D* produce a complicated interference pattern, which is recorded on light-sensitive film *F*. The structure thus obtained is called the reference-free hologram. Each point of an object (diffuser) recorded on such a hologram, e.g., point P_2, may be viewed as a coherent source relative to the remaining points of the object (P_1, P_3, …). The entire hologram *H* can be represented as a sum of subholograms sh_1, sh_2, sh_3, …:

$$H = sh_1 + sh_2 + sh_3 + \dots .$$

Each of the subholograms may be independently read by one of the points of the diffuser. In this case, this point can be viewed as a reference source for the given hologram.

The scheme for reconstructing image D_R of diffuser *D*, which serves as a visibility zone when 2D images are projected, is shown in Fig. 2. Hologram *H* with the image of the diffuser is illuminated by the radiation from point P_2', which coincides with one of the points (point P_2 in this case) of the image of diffuser *D* (Fig. 1) recorded on the hologram. Interacting with an appropriate subhologram, this radiation reconstructs virtual image D_V of diffuser *D*. Lens *L* behind hologram *H* converts virtual image D_V to real image D_R.

The scheme for 2D image projecting with such a screen is shown in Fig. 3. Projection lens L_p placed at the position of point P_2' (Fig. 1) projects image *I* of a scene onto screen *H*. Lens *L* behind the screen produces

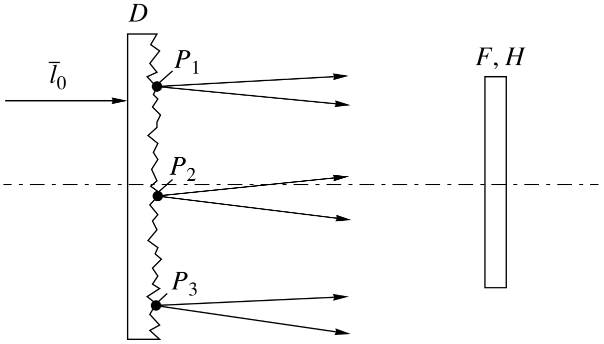


Fig. 1. Scheme for recording the diffuse screen. *D*, diffuser illuminated by coherent radiation $\bar{l}_0$; (P_1, P_2, P_3, …), separate points of diffuser *D*; and *F*, light-sensitive film recording hologram *H*.

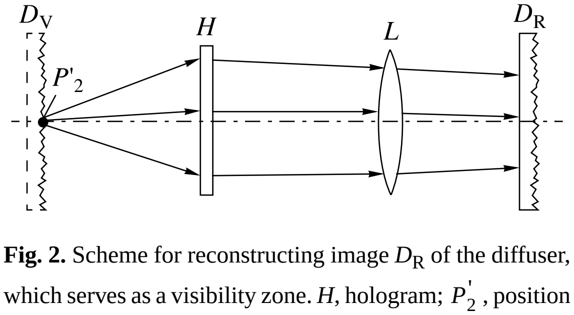


Fig. 2. Scheme for reconstructing image D_R of the diffuser, which serves as a visibility zone. *H*, hologram; P_2', position of the point reference source used to reconstruct the image of diffuser *D* recorded on the hologram; D_V, virtual image of diffuser *D*; and *L*, lens converting the virtual image D_V of diffuser *D* to its real image D_R.

real image D_R of diffuser D. Image D_R serves as visibility zone VZ, through which observer h sees the image I of the scene focused on screen H.

In the experiments, a thick (0.10–0.15 mm) layer of self-developing bichromated gelatin sensitive in the blue range [1–3] was used as a recording medium. Data recorded on this material are visualized (developed) directly during exposure and are fixed when kept in the dark for several days. Thus, recording of the diffuser hologram may be controlled immediately during exposure.

Figure 4 demonstrates the experimental setup used for recording and reconstruction of the screens. Lens L_0 focuses the laser radiation into point S and then produces an illuminated spot of diameter E_D = 10 mm on diffuser D. Diffuser D made by the special technology scatters the light within a small angle (about 0.03 rad). As a result, the diameter E_H of the illuminated spot on the surface of light-sensitive film F is about 20 mm. Using the directional diffuser, one can completely utilize the energy of the radiation exposing light-sensitive film F.

Specifically, early in the exposure, only the zeroth order (bright point P_2', which is the image of point P_2 produced by lens L) is seen at the center of the plane of visibility zone VZ (Fig. 4). As the exposure increases, the point image of the zeroth order diffuses and the visibility zone is uniformly filled with diffuse light. Point images related to higher diffraction orders were absent. So were a halo and the conjugate image of the diffuser.

The most interesting feature of the reference-free volume hologram obtained in our experiments is that, being recorded by the radiation with λ = 442 nm, it allows reconstruction by the radiation with λ = 633 nm. Such an effect, which comes into conflict with the selectivity rules the volume hologram obeys, may be explained by the fact that hologram H includes not only the subhologram corresponding to point P_2 but also many subholograms corresponding to other points of diffuser D.

The fact that the reference-free hologram can be reconstructed by radiation of a wavelength much different from the recording wavelength virtually means that the hologram may be recorded in the sensitivity range of the photographic film and then be used to project color images.

Thus, having taken the Gabor hologram as the basis, we devised a new type of hologram that is capable of compensating for the projected image distortions typical of the Gabor hologram. Specifically, based on the reference-free thick hologram, an on-axis holographic screen was developed that does not transmit the zeroth order. Nor does this screen produce a halo and the conjugate image. Also, it allows for recording in one spectral range and reconstruction in another. Such a feature simplifies the choice of a record medium and makes it possible to project a color image through the screen.

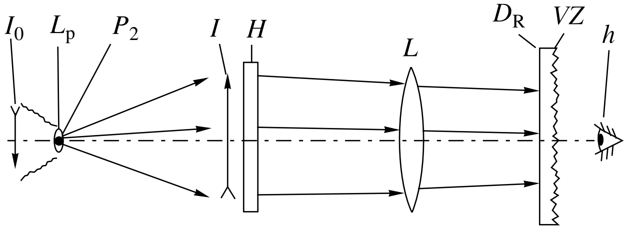


Fig. 3. Scheme for projecting 2D images by means of the screen developed. L_p, projection lens placed at the position of point P_2' (Fig. 1); L, lens producing real image D_R of diffuser D, which is visibility zone VZ for observer h; and I_0, image of scene I focused on screen H.

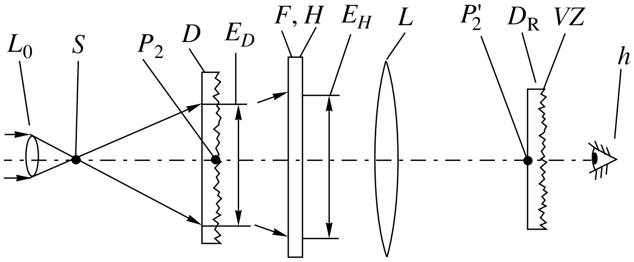


Fig. 4. Experimental setup used to record and reconstruct the on-axis screens. L_0, lens focusing the laser radiation into point S; D, highly directional diffuser with an illuminated spot of diameter E_D; F, light-sensitive film in the plane of which the diffuse light spot has a diameter E_H; L, lens producing the real image D_R of the diffuser during reconstruction (here, the diffuser serves as visibility zone VZ for observer h).

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 04-02-17593) and grant no. NSh-98.2003.2 of the President of the Russian Federation (“Program of Support of Leading Scientific Schools”).

REFERENCES

1. Yu. N. Denisyuk, N. M. Ganzherli, I. A. Maurer, and S. A. Pisarevskaya, Pis’ma Zh. Tekh. Fiz. **23** (4), 62 (1997) [Tech. Phys. Lett. **23**, 279 (1997)].
2. Yu. N. Denisyuk, N. M. Ganzherli, I. A. Maurer, and S. A. Pisarevskaya, Proc. SPIE **3011**, 244 (1997).
3. Yu. N. Denisyuk, N. M. Ganzherli, I. A. Maurer, and S. A. Pisarevskaya, Pis’ma Zh. Tekh. Fiz. **25** (5), 64 (1999) [Tech. Phys. Lett. **25**, 194 (1999)].

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 276–279. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 137–139.
Original Russian Text Copyright © 2005 by Fedorov, Epshtein, Shmelev.

BRIEF COMMUNICATIONS

Quasi-One-Dimensional Ballistic Ring in Crossed RF Electric Fields

E. G. Fedorov, E. M. Epshtein, and G. M. Shmelev
Volgograd State University of Architecture and Civil Engineering, Volgograd, 400074 Russia
e-mail: eduard-f@mail.ru

Received June 7, 2004

Abstract—The electron dynamics in a quasi-one-dimensional ballistic ring that is subjected to two rf electric fields that have mutually orthogonal polarizations and lie in the plane of the ring is considered. The mean dipole moment and the ring radiation intensity are calculated. The condition for spontaneous symmetry breaking, as a result of which the dipole moment of the system acquires a constant component, is found. *© 2005 Pleiades Publishing, Inc.*

Advances in the technology of mesoscopic structures are encouraging experimental and theoretical investigation into low-dimensional one-electron systems, specifically, quasi-one-dimensional rings (see, e.g., [1, 2]). Most of the related works were devoted to quantum phenomena. It has been shown [3–9], however, that quasi-one-dimensional rings also offer intriguing classical electrodynamic properties. In this work, we study the response of an electron placed into a quasi-one-dimensional ring to two external electromagnetic (EM) waves with mutually orthogonal polarizations that propagate along the normal to the plane of the ring (the electric vectors of the waves lie in the plane of the ring).

Consider a planar ring of radius R (the thickness of the ring is small compared with its radius) that represents a quantum well between two concentric potential barriers. Along the radius, the electron is assumed to execute quantum motion; along the circumference, its motion is classical. We also assume that the free path of the electron is much larger than $2\pi R$ (ballistic motion) and the wavelengths far exceed the diameter of the ring. Then, in the dipole approximation, the electron experiences the action of only the electric fields of the waves,

$$\mathbf{E} = \{E_1\sin(\omega_1 t + \beta_1),\ E_2\sin(\omega_2 t + \beta_2)\}. \tag{1}$$

It is assumed that, at the zero time $t = 0$ (the fields are absent), the electron has energy W and circulates along the circumference. The equation of motion of the electron in the ring has the form

$$\begin{aligned}&\ddot{\varphi} + \Omega_1^2\sin\varphi\sin(\omega_1 t + \alpha_1)\\ &-\Omega_2^2\cos\varphi\sin(\omega_2 t + \alpha_2) = 0,\end{aligned} \tag{2}$$

where $\Omega_{1,2}^2 = |e|E_{1,2}/(mR)$, e and m are the charge and effective mass of the electron, $\alpha_{1,2} = \beta_{1,2} \pm \pi$, and φ is the angular coordinate measured from the Ox axis (which is parallel to field $\mathbf{E}_1$).

Field $\mathbf{E}$ is assumed to be radio-frequency: $\omega_{1,2} \gg \omega_{\text{slow}}$ (ω_{slow} is the frequency of the "slow" component of the electron motion, see below). To gain insight into the slow motion of the electron, we will make use of the Kapitsa averaging method [10], which assumes that $\varphi = \Phi + \xi$, where Φ and ξ characterize the slow motion and high-frequency oscillations, respectively. The slow motion obeys the equation $mR^2\ddot{\Phi} = -\partial U_{\text{eff}}/\partial\Phi$, where

$$\begin{aligned}U_{\text{eff}} = mR^2\Bigg\{&\frac{1}{4}\left[\left(\frac{\Omega_1^2}{\omega_1}\right)^2 - \left(\frac{\Omega_2^2}{\omega_2}\right)^2\right]\sin^2\Phi\\ &+ \frac{\Omega_1^2\Omega_2^2}{2}A\left[\Phi\left(\frac{1}{\omega_1^2} - \frac{1}{\omega_2^2}\right) - \frac{1}{2}\left(\frac{1}{\omega_1^2} + \frac{1}{\omega_2^2}\right)\sin 2\Phi\right]\Bigg\}\end{aligned} \tag{3}$$

is the effective potential energy and $A = 0$ if $\omega_1 \neq \omega_2$ or $A = (1/2)\cos(\alpha_1 - \alpha_2)$ if $\omega_1 = \omega_2$.

Let us consider a number of special cases following from expression (3).

(1) If $\omega_1 \neq \omega_2$, we get from (3)

$$U_{\text{eff}} = \frac{mR^2}{4}\left[\left(\frac{\Omega_1^2}{\omega_1}\right)^2 - \left(\frac{\Omega_2^2}{\omega_2}\right)^2\right]\sin^2\Phi. \tag{4}$$

If $\Omega_1^2/\omega_1 > \Omega_2^2/\omega_2$, potential energy (4) has minima at $\Phi = 0$ and $\Phi = \pi$. If $\Omega_1^2/\omega_1 < \Omega_2^2/\omega_2$, the potential energy is minimal at $\Phi = \pm\pi/2$.

(2) If $\omega_1 = \omega_2 = \omega$, we have from (3)

$$\begin{aligned}U_{\text{eff}} = \frac{mR^2}{4\omega^2}[&(\Omega_1^4 - \Omega_2^4)\sin^2\Phi\\ &-\Omega_1^2\Omega_2^2\cos(\alpha_1 - \alpha_2)\sin 2\Phi],\end{aligned} \tag{5}$$

1063-7842/05/5002-0276$26.00 © 2005 Pleiades Publishing, Inc.

which yields

$$U_{\text{eff}} = -\frac{mR^2}{4\omega^2}\Omega^4\cos(\alpha_1-\alpha_2)\sin 2\Phi \qquad (6)$$

if $\Omega_1 = \Omega_2 = \Omega$.

If $\alpha_1 - \alpha_2 \neq \pm\pi/2$, potential energy (6) has minima at $\Phi = \pi/4$ and $\Phi = 5\pi/4$.

(3) If $\omega_1 = \omega_2 = \omega$ and $\alpha_1 - \alpha_2 = \pm\pi/2$ (elliptically polarized EM wave), expression (3) takes the form

$$U_{\text{eff}} = \frac{mR^2}{4\omega^2}(\Omega_1^4 - \Omega_2^4)\sin^2\Phi. \qquad (7)$$

(4) If $\Omega_1^2/\omega_1 = \Omega_2^2/\omega_2$, it follows from (3) that $U_{\text{eff}} = 0$ at $\omega_1 \neq \omega_2$ and/or $\alpha_1 - \alpha_2 = \pm\pi/2$.

From expression (4) for the potential energy, we obtain an equation for the slow (averaged over rf oscillations) electron motion:

$$\ddot{\Phi} + \frac{1}{4}\left[\left(\frac{\Omega_1^2}{\omega_1}\right)^2 - \left(\frac{\Omega_2^2}{\omega_2}\right)^2\right]\sin 2\Phi = 0. \qquad (8)$$

A solution to equation of motion (8) subject to the initial conditions $\Phi_0 = \Phi(0)$ and $\dot{\Phi}_0 = \dot{\Phi}(0) = R^{-1}(2W/m)^{1/2}$ has the form

$$\Phi = \Psi + \lambda, \qquad (9)$$

where

$$\Psi = \begin{cases} \arcsin\{q\,\mathrm{sn}[\omega_0 t \\ \quad + F(\arcsin(q^{-1}\sin(\Phi_0-\lambda)), q), q]\}, & q < 1, \\ \arcsin\{\tanh[\omega_0 t + F(\Phi_0 - \lambda, 1)]\}, & q = 1, \\ \arcsin\{\mathrm{sn}[q\omega_0 t + F(\Phi_0 - \lambda, q^{-1}), q^{-1}]\}, & q > 1, \end{cases} \qquad (10)$$

$\lambda = 0$ if $\Omega_1^2/\omega_1 > \Omega_2^2/\omega_2$ or $\lambda = \pi/2$ if $\Omega_1^2/\omega_1 < \Omega_2^2/\omega_2$,

$$\omega_0 = \sqrt{\frac{1}{2}\left|\left(\frac{\Omega_1^2}{\omega_1}\right)^2 - \left(\frac{\Omega_2^2}{\omega_2}\right)^2\right|}, \qquad (11)$$

$$q^2 = \frac{\dot{\Phi}_0^2}{\omega_0^2} + \frac{1}{2}(1 - \cos 2(\Phi_0 - \lambda)), \qquad (12)$$

$\mathrm{sn}(\chi, k)$ is the elliptic sine, and $F(\chi, k)$ is the elliptic integral of the first kind [11].

The fundamental frequency of the slow motion of the electron is given by

$$\omega_{\text{slow}} = \omega_0\frac{\pi}{2}\begin{cases} \dfrac{1}{K(q)}, & q < 1 \\ 0, & q = 1 \\ \dfrac{q}{K(q^{-1})}, & q > 1, \end{cases} \qquad (13)$$

where $K(\chi)$ is the complete elliptic integral of the first kind [11].

The specific (per electron) dipole moment (relative to the center of the ring) and the ring radiation intensity are expressed as [12]

$$\mathbf{P} = eR\{\cos\Phi, \sin\Phi\}, \qquad (14)$$

$$I = \frac{2}{3c^3}\ddot{\mathbf{P}}^2 = \frac{2}{3c^3}(eR)^2(\dot{\Phi}^4 + \ddot{\Phi}^2). \qquad (15)$$

The ring radiation intensity is found by substituting (9) into (15),

$$I = \frac{2}{3c^3}(eR)^2\omega_0^4$$

$$\times\begin{cases} q^4 + q^2(1-2q^2)\mathrm{sn}^2[\omega_0 t \\ \quad + F(\arcsin(q^{-1}\sin(\Phi_0-\lambda)), q), q], & q < 1, \\ \mathrm{sech}^2[\omega_0 t + F(\Phi_0 - \lambda, 1)], & q = 1, \\ q^4 + (1-2q^2)\mathrm{sn}^2[q\omega_0 t \\ \quad + F(\Phi_0 - \lambda, q^{-1}), q^{-1}], & q > 1. \end{cases} \qquad (16)$$

The mean radiation intensity is given by

$$\langle I\rangle = \frac{2}{3c^3}(eR)^2\omega_0^4 G(q), \qquad (17)$$

where function $G(q)$ has the form

$$G(q) = \begin{cases} q^4 + (1-2q^2)\left[1 - \dfrac{E(q)}{K(q)}\right], & q < 1 \\ 0, & q = 1 \\ q^4 + q^2(1-2q^2)\left[1 - \dfrac{E(q^{-1})}{K(q^{-1})}\right], & q > 1 \end{cases} \qquad (18)$$

and $E(\chi)$ is the complete elliptic integral of the second kind [11].

At $q \ll 1$, the electron weakly oscillates and $G(q) \cong q^2/2$; then, (17) yields

$$\langle I\rangle = \frac{1}{3c^3}(eR)^2\omega_0^4 q^2. \qquad (19)$$

At $q \gg 1$ (i.e., $\omega_0^2 \ll \dot{\Phi}_0^2$), we have $q^2 \simeq \dot{\Phi}_0^2/\omega_0^2$; therefore, $G(q) \cong q^4$, $(\omega_0 q)^4 \cong \dot{\Phi}_0^4$, and (17) yields

$$\langle I\rangle = I_0 = \frac{2}{3c^3}(eR)^2\dot{\Phi}_0^4. \qquad (20)$$

Expression (20) coincides with the expression for the emission of an electron describing a circle of radius

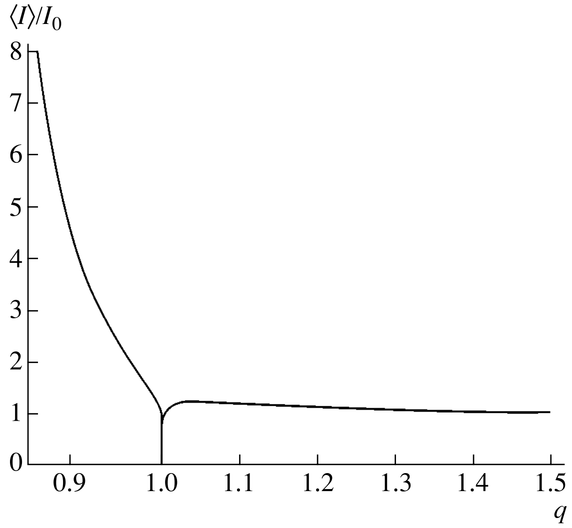


Fig. 1. $\langle I\rangle/I_0$ vs. $q = q(\dot{\Phi}_0^2/\omega_0^2)$ at $\Phi_0 = \pi/4$.

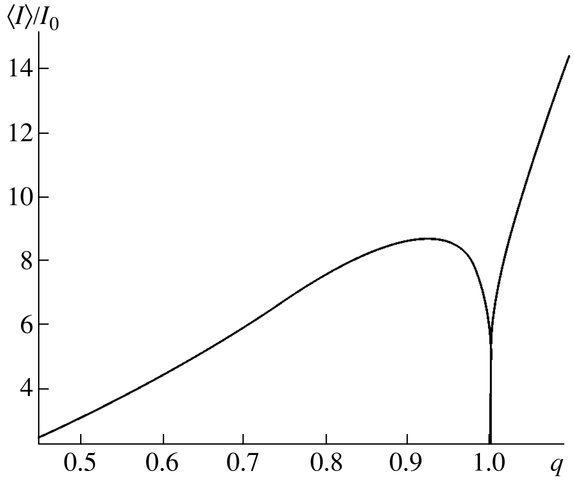


Fig. 2. $\langle I\rangle/I_0$ vs. $q = q(\Phi_0)$ at $\dot{\Phi}_0/\omega_0 = 0.45$.

R with a constant angular velocity [12] that corresponds to energy $W = R^2\dot{\Phi}_0^2 m/2$.

Consider function $\langle I\rangle(q)$ where q is a function of either $\dot{\Phi}_0^2/\omega_0^2$ or Φ_0:

$$\langle I\rangle = I_0\frac{G(q)}{q^4}\left[1 + M(q) + \frac{1}{4}M^2(q)\right], \quad (\Phi_0 = fix), \tag{21}$$

where

$$M(q) = \frac{(1-\cos 2(\Phi_0-\lambda))}{q^2-(1/2)(1-\cos 2(\Phi_0-\lambda))}$$

or

$$\langle I\rangle = I_0\frac{G(q)}{q^4}\left[1 + B(q) + \frac{1}{4}B^2(q)\right], \quad (\dot{\Phi}_0^2/\omega_0^2 = fix), \tag{22}$$

where $B(q) = 2(q^2\omega_0^2/\dot{\Phi}_0^2 - 1)$.

Figures 1 and 2 plot the curves $\langle I\rangle/I_0(q)$ for $\Omega_1^2/\omega_1 > \Omega_2^2/\omega_2$ that are constructed by formulas (21) and (22), respectively.

The mean dipole moment $\langle\mathbf{P}\rangle$ of the ring is found by substituting (9) into (14) and averaging the result over period $T_{\text{slow}} = 2\pi/\omega_{\text{slow}}$. At $q > 1$, we have $\langle\mathbf{P}\rangle = \{0, 0\}$. At $q < 1$, two cases are possible:

$$\langle\mathbf{P}\rangle = eR\left\{\pm\frac{\pi}{2K(q)}, 0\right\}$$

for $\Omega_1^2/\omega_1 > \Omega_2^2/\omega_2$ and

$$\langle\mathbf{P}\rangle = eR\left\{0, \pm\frac{\pi}{2K(q)}\right\}$$

for $\Omega_1^2/\omega_1 < \Omega_2^2/\omega_2$. This effect can be explained as follows. When exposed to the electric field of an incident EM wave, the slow motion of the electron is motion inside a well in the two-well potential relief (see (4)). At $q > 1$, the energy of the electron exceeds the height of the potential barrier between the two wells and the electron overcomes the barrier. At $q < 1$, the electron oscillates within one well. Thus, at $q = 1$, the symmetry spontaneously breaks and the dipole moment of the system may acquire a constant component (optical rectification) [3].

If the ring contains N electrons, the expressions for $\mathbf{P}$ and $\langle\mathbf{P}\rangle$ are multiplied by N and those for I and $\langle I\rangle$, by N^2 (provided that Φ_0 and $\dot{\Phi}_0$ are the same for all electrons).

Let us make numerical estimates. For $R = 5 \times 10^{-5}$ cm, $m = 0.1m_e$ (m_e is the mass of a free electron), $W = 2 \times 10^{-3}$ eV, $\omega_1 = 10^{13}$ s^{-1}, $\omega_2 = 10^{14}$ s^{-1}, $E_1 = 1.5 \times 10^4$ V/cm, and $E_2 = 3 \times 10^3$ V/cm, we have $\Omega_1 \approx 2.30 \times 10^{12}$ s^{-1}, $\Omega_2 \approx 1.03 \times 10^{12}$ s^{-1}, $\omega_0 \approx 3.73 \times 10^{11}$ s^{-1}, $\dot{\Phi}_0/\omega_0 \approx 0.45$, and $\Omega_1^2/\omega_1 > \Omega_2^2/\omega_2$. For $\Phi_0 = \pi/4$ and the parameter values listed above, we obtain $q \approx 0.84$ (see (12)) and $\omega_{\text{slow}} \approx 2.82 \times 10^{11}$ s^{-1} (see (13)). Thus, it follows from (21) and (22) that $\langle I\rangle/I_0 \approx 8.06$. For the EM wave frequencies used in this work, the inequality $\omega_{1,2} \gg \omega_{\text{slow}}$, which is the condition necessary for the Kapitsa averaging method to apply, holds. For the electron collision frequency $\nu = 5 \times 10^9$ s^{-1}, we have $2\pi\nu \ll$

ω_{slow}, which means that the ballistic approximation works well in description of the electron dynamics in the ring.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research, grant no. 02-02-16238.

REFERENCES

1. A. Lorke, R. J. Luyken, A. O. Govorov, *et al.*, Phys. Rev. Lett. **84**, 2223 (2000).
2. N. T. Bagraev, A. D. Buravlev, V. K. Ivanov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 846 (2000) [Semiconductors **34**, 817 (2000)].
3. E. M. Epshtein, G. M. Shmelev, and I. I. Maglevanny, J. Phys. A **33**, 6017 (2000).
4. E. M. Epshtein and G. M. Shmelev, Phys. Scr. **62**, 216 (2000).
5. G. M. Shmelev, E. M. Epshtein, and G. A. Syrodoev, Zh. Tekh. Fiz. **70** (10), 125 (2000) [Tech. Phys. **45**, 1354 (2000)].
6. E. M. Epshtein, I. I. Maglevanny, and G. M. Shmelev, Phys. Low-Dimens. Semicond. Struct. **3/4**, 109 (2000).
7. E. M. Epshtein, G. M. Shmelev, and I. I. Maglevanny, Phys. Low-Dimens. Semicond. Struct. **1/2**, 137 (2001).
8. G. M. Shmelev and E. M. Epshtein, Fiz. Tverd. Tela (St. Petersburg) **43**, 2213 (2001) [Phys. Solid State **43**, 2311 (2001)].
9. E. M. Epshtein and G. M. Shmelev, cond-mat/0312025.
10. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics* (Nauka, Moscow, 1982; Pergamon, New York, 1988).
11. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1978).
12. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, pp. 280–283. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, pp. 140–142.
Original Russian Text Copyright © 2005 by Lomtev.

Surface Polaritons at the Interface between a Polarized Dirac Vacuum and a Normal Metal

A. I. Lomtev
Galkin Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 83114 Ukraine
e-mail: lomtev@kinetic.ac.donetsk.ua

Received June 23, 2004

Abstract—The behavior of surface polaritons that can propagate along a flat interface between a polarized Dirac vacuum and a normal metal is studied. Dispersion relations for the surface wave are derived for different field polarizations, and the frequency domains where this wave may exist are found. It is shown that a constant electric field polarizing a Dirac vacuum can be determined from experiments on surface wave excitation. © *2005 Pleiades Publishing, Inc.*

Electromagnetic waves propagating along flat interfaces between materials with different insulating properties continue to be a subject of keen interest [1–3]. It is known that a surface *H* wave (*P* polarization) may propagate along the interface between media with positive and negative permittivities (ε_1 and $-|\varepsilon_2|$, respectively), which decays on both sides of the interface [4].

The term *surface wave* was first introduced in the context of the theory of rf wave propagation over the terrestrial surface. As early as in 1907, Zenneck [5, 6] theoretically showed that a slow rf wave may travel over conducting areas of the terrestrial (or sea) surface. Basically, this wave is similar to that propagating along a conductor.

The optical properties of free electron bunches [7; Sects. 13, 48] allow investigation of their internal structure. According to Dirac [8], vacuum may be conceived as a set of free electrons occupying negative-energy levels. It is shown in this work that, considering surface polaritons at the polarized Dirac vacuum–normal metal interface, one can study very high constant electric fields $E_0 < E_{\mathrm{cr}} = m^2c^3/e\hbar = 1.3 \times 10^{16}$ V/cm (otherwise, i.e., for $E_0 > E_{\mathrm{cr}}$, breakdown of the vacuum will take place).

A Dirac vacuum polarized by field $\mathbf{E}_0$ exhibits not only insulating but also magnetic properties. Its permittivity and permeability tensors have the form [9]

$$\varepsilon_{1ij} = (1 + 8\pi R E_0^2)\delta_{ij} + 16\pi R E_{0i}E_{0j}, \tag{1}$$

$$\mu_{1ij} = (1 - 8\pi R E_0^2)\delta_{ij} + 56\pi R E_{0i}E_{0j}, \tag{2}$$

where $R = \alpha/360\pi^2 E_{\mathrm{cr}}^2$ and $\alpha = e^2/\hbar c$.

Note that formulas (1) and (2) apply to both constant and variable fields if the frequency of the latter is $\omega \ll mc^2/\hbar$ [10].

The permittivity of a normal metal is given by [11, 12]

$$\varepsilon_{2ij} = \varepsilon_2\delta_{ij} = [1 - \omega_{\mathrm{p}}^2/\omega(\omega + i/\tau)]\delta_{ij}, \tag{3}$$

where $\omega_{\mathrm{p}} = (4\pi n e^2/m^*)^{1/2}$ is the electron plasma frequency, n is the electron density, e is the electron charge, m^* is the effective mass of an electron, and τ is the electron relaxation time (the time between collision of electrons with defects or impurities).

Formula (3) for the permittivity describes the contribution of intraband transitions to the permittivity of a gas of near-free electrons in a metal or *n*-semiconductor. In high-purity metals, $\tau \sim 10^{-9}$ s; for typical metals, $\omega_{\mathrm{p}} \sim 10^{16}$ s^{-1} [12]. Therefore, in the range of positive dispersion $\tau^{-1} \ll \omega < \omega_p$, where $\omega = 10^{15}$–10^{16} s^{-1}, the imaginary part of permittivity (3) may be safely neglected and the decay may be considered weak. Then, permittivity (3) becomes a real and negative function of frequency,

$$\varepsilon_{2ij} = \varepsilon_2\delta_{ij} = -(\omega_{\mathrm{p}}^2/\omega^2 - 1)\delta_{ij}, \tag{4}$$

and the metal represents a surface-active medium [13].

As a metal, it is convenient to take an alkali metal, e.g., Na, with a spherical Fermi surface, which provides the isotropy of its permittivity. The permeability of a nonmagnetic metal is $\mu_{2ij} = \delta_{ij}$.

Consider the propagation of a surface electromagnetic *H* wave along the polarized Dirac vacuum–normal metal interface. Let the interface be plane *xy*. We assume that the wave propagates along the *x* axis and field $\mathbf{H}$ is directed along the *y* axis. For *P*-polarized waves, which are proportional to $\exp(-i\omega t)$, with nonzero components $\{E_x, H_y, E_z\}$, the Maxwell equations have the form

$$\partial H_y/\partial z = i(\omega/c)D_x, \quad \partial H_y/\partial x = -i(\omega/c)D_z, \quad \partial E_x/\partial z - \partial E_z/\partial x = i(\omega/c)B_y, \tag{5}$$

where $\mathbf{D} = \tilde{\varepsilon}\,\mathbf{E}$ and $\mathbf{B} = \tilde{\mu}\,\mathbf{H}$.

Let the half-space $z > 0$ be filled with a polarized Dirac vacuum with a positive ($\tilde{\varepsilon}_1$) permittivity and the half-space $z < 0$, with a normal metal with a negative ($\tilde{\varepsilon}_2$) permittivity. The field in the wave decaying at infinity ($z \longrightarrow \pm\infty$) is sought in the form

$$H_1 = H_0\exp(ikx - \kappa_1 z) \quad \text{at} \quad z > 0, \tag{6}$$

$$H_2 = H_0\exp(ikx + \kappa_2 z) \quad \text{at} \quad z < 0, \tag{7}$$

where k, κ_1, and κ_2 are real. The boundary condition $H_1(z = 0) = H_2(z = 0)$ has been already satisfied, and the continuity condition for E_x yields a dispersion relation $k = k(\omega)$,

$$\frac{1}{\varepsilon_{1xx}}\frac{\partial H_1}{\partial z} = \frac{1}{\varepsilon_2}\frac{\partial H_2}{\partial z} \quad \text{at} \quad z = 0 \tag{8}$$

or

$$\frac{\kappa_1}{\varepsilon_{1xx}} = \frac{\kappa_2}{|\varepsilon_2|}. \tag{9}$$

It will be seen below that $\kappa_2 > \kappa_1$. Then, the inequality $|\varepsilon_2| > \varepsilon_{1xx}$ is a necessary condition for the surface waves to exist. This inequality imposes an upper limit on the allowable frequency range: $\omega < \omega_p/(1 + \varepsilon_{1xx})^{1/2}$. Consequently, the surface wave frequency must satisfy the inequalities $\tau^{-1} \ll \omega < \omega_p/(1 + \varepsilon_{1xx})^{1/2}$.

From Maxwell equations (5) for field (7), it follows that

$$\kappa_2^2 = k^2 + \frac{\omega_\mathrm{p}^2 - \omega^2}{c^2}. \tag{10}$$

Now let us assume that $\mathbf{E}_0$ (polarizing field) = (E_{0x}, 0, 0). According to Maxwell equations (5) for field (6), an expression for κ_1 for this direction of the polarizing field takes the form

$$\kappa_1^2 = (1 + 24\pi RE_{0x}^2) \times \left[\frac{k^2}{1 + 8\pi RE_{0x}^2} - \frac{\omega^2}{c^2}(1 - 8\pi RE_{0x}^2)\right]. \tag{11}$$

In view of (10) and (11), dispersion relation (9) is expressed as

$$k^2(\omega) = \frac{\omega^2}{c^2} \times \left\{\frac{(1 + 8\pi RE_{0x}^2)(\omega_\mathrm{p}^2 - \omega^2)[(1 - 8\pi RE_{0x}^2)\omega_\mathrm{p}^2 + 32\pi RE_{0x}^2\omega^2]}{\omega_\mathrm{p}^2(\omega_\mathrm{p}^2 - 2\omega^2) - 32\pi RE_{0x}^2\omega^4(1 + 6\pi RE_{0x}^2)}\right\}. \tag{12}$$

Then, the surface wave frequency must satisfy the inequalities

$$\tau^{-1} \ll \omega < \omega_\mathrm{p}/\sqrt{2(1 + 12\pi RE_{0x}^2)}.$$

If the polarizing field is aligned with the y axis, $\mathbf{E}_0$ = (0, E_{0y}, 0), κ_1 appearing in field (6) is given, according to (5), by

$$\kappa_1^2 = (1 + 8\pi RE_{0y}^2) \times \left[\frac{k^2}{1 + 8\pi RE_{0y}^2} - \frac{\omega^2}{c^2}(1 + 48\pi RE_{0y}^2)\right], \tag{13}$$

and, in view of (10), we come to the dispersion relation

$$k^2(\omega) = \frac{\omega^2}{c^2} \times \left\{\frac{(1 + 8\pi RE_{0y}^2)(\omega_\mathrm{p}^2 - \omega^2)[(1 + 48\pi RE_{0y}^2)\omega_\mathrm{p}^2 - 40\pi RE_{0y}^2\omega^2]}{\omega_\mathrm{p}^2(\omega_\mathrm{p}^2 - 2\omega^2) - 16\pi RE_{0y}^2\omega^4(1 + 4\pi RE_{0y}^2)}\right\}. \tag{14}$$

In this case, the surface wave frequency must satisfy the inequalities

$$\tau^{-1} \ll \omega < \omega_\mathrm{p}/\sqrt{2(1 + 4\pi RE_{0y}^2)}.$$

If the polarizing field is aligned with the z axis, $\mathbf{E}_0$ = (0, 0, E_{0z}), κ_1 appearing in field (6) is given, according to (5), by

$$\kappa_1^2 = (1 + 8\pi RE_{0z}^2) \times \left[\frac{k^2}{1 + 24\pi RE_{0z}^2} - \frac{\omega^2}{c^2}(1 - 8\pi RE_{0z}^2)\right] \tag{15}$$

and, in view of (10), we come to the dispersion relation

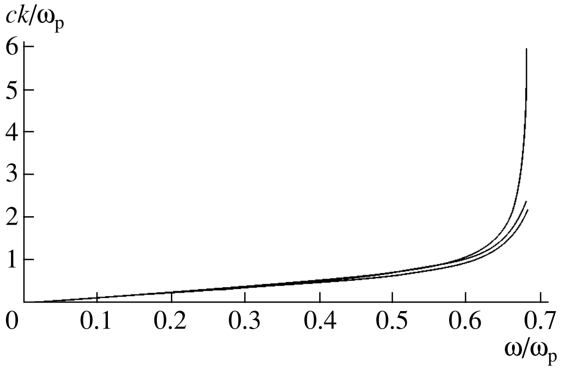
$$k^2(\omega) = \frac{\omega^2}{c^2} \times \left\{ \frac{(1 + 24\pi RE_{0z}^2)(\omega_p^2 - \omega^2)[(1 - 8\pi RE_{0z}^2)\omega_p^2 + 16\pi RE_{0z}^2\omega^2]}{\omega_p^2(\omega_p^2 - 2\omega^2) - 32\pi RE_{0z}^2\omega^4(1 + 6\pi RE_{0z}^2)} \right\}. \tag{16}$$

In this case, the surface wave frequency must satisfy the inequalities

$$\tau^{-1} \ll \omega < \omega_p / \sqrt{2(1 + 4\pi RE_{0z}^2)}.$$

It is seen from dispersion relations (12), (14), and (16) that the field polarizing a Dirac vacuum causes the dependence of the surface wave frequency spectrum on its polarization.

The figure shows the dispersion curves for the surface polaritons that are obtained by numerically solving dispersion relations (12), (14), and (16). The difference between the spectrum branches, which is due to the different directions of the field polarizing the Dirac vacuum, is the most pronounced at frequencies close to $\omega = \omega_p/(1 + \varepsilon_{1xx})^{1/2}$. From expression (10) for κ_2 and dispersion relations (12), (14), and (16), it follows that the penetration depth of the electromagnetic field of the surface wave into the metal is $\delta = \kappa_2^{-1} \approx c/\omega_p = \sqrt{3}\, c/v_F k_s \gg r_{TF}$, where $v_F = 10^5$–10^6 cm/s is the Fermi velocity of electrons and $k_s^{-1} = r_{TF}$ is the Thomas–Fermi radius. On the other hand, the free path of an electron in a metal is $l = \tau v_F = \sqrt{3}\,\tau\omega_p r_{TF} \gg r_{TF}$. The ratio $\delta/l = (c/v_F)/\omega_p\tau$ is on the order of 10^{-3}–10^{-2}. Since $\delta \ll 1$, it may be considered that a surface wave in a metal experiences the anomalous skin effect.



Wavevector ck/ω_p vs. reduced frequency ω/ω_p for surface polaritons. The bottom curve is the solution to dispersion relation (12); the middle curve, to dispersion relation (14); and the top curve, to dispersion relation (16). The parameter values are $\varepsilon_{1xx} \approx 1$, $RE_{0x}^2 = 10^{-3}$, $RE_{0y}^2 = 2 \times 10^{-3}$, and $RE_{0z}^2 = 3 \times 10^{-3}$. The allowable frequency range is $10^{-5} \le \omega/\omega_p \le 0.7$.

The waves with dispersion relations (12), (14), and (16) are Zenneck surface waves [5, 6] propagating along the flat polarized Dirac vacuum–normal metal interface with the frequencies in the visible range.

If the volume of a polarized Dirac vacuum is experimentally accessible, the contribution of the polarizing field is the easiest to determine by measuring either the optical birefringence for linearly polarized light or the angle of rotation of polarization ellipse for elliptically polarized light [9]. If only the surface of a Dirac vacuum is accessible in an experiment, one readily finds the contributions RE_{0x}^2, RE_{0y}^2, or RE_{0z}^2 to the permittivity and permeability of a polarized Dirac vacuum from dispersion relations (12), (14), or (16) by optically exciting surface waves.

It is easy to check that the propagation of surface *E* waves (*S* polarization) with components $\{H_x, E_y, H_z\}$ along such an interface is impossible.

ACKNOWLEDGMENTS

The author thanks S.V. Tarasenko for valuable discussion and also Yu.V. Medvedev and I.B. Krasnyuk for support.

REFERENCES

1. V. L. Falko, S. I. Khankina, and V. V. Yakovenko, Phys. Lett. A **209**, 118 (1995).
2. M. F. Pascual, W. Zierau, T. F. Leskova, and A. A. Maradudin, Opt. Commun. **155**, 351 (1998).
3. N. L. Dmitruk, S. V. Mamykin, and O. V. Rengevych, Appl. Surf. Sci. **166** (1–4), 97 (2000).
4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
5. J. Zenneck, Ann. Phys. **3**, 846 (1907).
6. L. A. Vainshtein, *Electromagnetic Waves* (Sov. Radio, Moscow, 1957) [in Russian].

7. V. M. Harutunian and S. G. Oganesyan, Phys. Rep. **270**, 217 (1976).
8. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Relativistic Quantum Theory*, 2nd ed. (Nauka, Moscow, 1971; Pergamon, Oxford, 1971), Part 2.
9. S. G. Oganesyan, Pis'ma Zh. Tekh. Fiz. **24** (14), 45 (1998) [Tech. Phys. Lett. **24**, 561 (1998)].
10. A. A. Grib, S. G. Mamaev, and V. M. Mostepanenko, *Vacuum Quantum Effects in Strong Fields* (Nauka, Moscow, 1988) [in Russian].
11. D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963; Mir, Moscow, 1965).
12. P. M. Platzman and P. A. Wolff, in *Solid State Physics* (Academic, New York, 1973; Mir, Moscow, 1975), Suppl. 13.
13. *Surface Polaritons*, Ed. by V. M. Agranovich and D. L. Mills (North-Holland, Amsterdam, 1982; Nauka, Moscow, 1985).

Translated by V. Isaakyan

Technical Physics, Vol. 50, No. 2, 2005, p. 284. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 2, 2005, p. 143.
Original Russian Text Copyright © 2005 by the Editorial Board.

OBITUARY

In Memory of Sergeĭ A. Kornilov

S.A. Kornilov, Dr. Sci., Prof., Honored Scientist of the Russian Federation, and a member of the editorial board of the *Tekhnicheskaya Fizika* journal, died October 27, 2004, after a serious disease.

Kornilov was born July 4, 1927, in Tambov in the family of a serviceman. In 1938, his father was subjected to repression, and Sergeĭ and his mother left Tambov for Leningrad. In spring 1942, being seriously ill after the blockade winter, Kornilov was conveyed to the "continent" on the last trip. After the war, he entered the Moscow Institute of Steel and Alloys. However, dreaming of the sea from childhood, he made a decision to leave the Moscow institute for the Leningrad Ship-Building Institute but was not admitted because of his bad eyesight. He entered the Leningrad Polytechnical Institute, Department of Physics and Mathematics, and graduated from that institute in 1950. In 1954, he defended his Cand. Sci. dissertation and, in 1968, doctoral dissertation. In 1974, Kornilov became the head of the Chair of Radio Engineering and, in 1975, the dean of the Department of Radio Physics. From 1978 to 1993, he worked as the head of the Chair of Electron and Quantum Devices at the Bonch-Bruevich Electrotechnical Institute of Communication. In the last few years, he worked in this chair as a professor.

Kornilov is the author of 120 scientific publications, including 2 monographs. Forty disciples of Kornilov have defended Cand. Sci. dissertations, and two of them doctoral dissertations.

The scientific activity of Kornilov was intimately related to industry. In 1958, he founded a laboratory of microwave electronics at the Chair of Radio Engineering, Leningrad Polytechnical Institute, which met the orders of the Ministry of Electronic Industry of the Soviet Union and became the forge of brainpower for this branch of industry. Initially, the laboratory designed and created microwave tubes. The unique K-11 frequency-multiplier klystron, with a frequency multiplication factor of 11, was produced by the Svetlana factory for more than 30 years.

From 1960 to his last days, the basic field of Kornilov's scientific interest was the physics of noise and fluctuation in electron tubes and solid-state electron devices. A specific feature of the scientific school founded by Kornilov was unity of theory and experiment. He developed the quasi-static method of fluctuation analysis for microwave devices, as well as the adequate noise models of amplifier, multiplier, and reflex klystrons; magnetrons and amplitrons; and oscillators and amplifiers based on tunnel diodes, avalanche diodes, bipolar and field-effect transistors, and Gunn diodes. In addition, Kornilov developed the noise models of microwave mixers and detectors based on Schottky-barrier diodes. Finally, he devised high-precision techniques for noise and fluctuation measurement, which are used in the industry as state standards, specifically, by manufacturers of high-sensitive and ultra-low-noise microwave systems.

Kornilov was a highly educated person. He knew German and English. At the age of forty, he started studying Spanish and leaned this language perfectly. He was enthusiastic about poetry and began to translate Spanish poems. Literary men consider Kornilov's translations of M. Lorca and O. Pasa into Russian as the best of those currently available. Some of the translations were published in the *Vsemirnaya Literatura* journal. A small collection of the Lorca poems translated by Kornilov was issued by the *Boreĭ* publishing house.

Those who knew Sergeĭ A. Kornilov will remember him as a man of mighty intellect and high moral principles.

Editorial Board

1063-7842/05/5002-0284$26.00 © 2005 Pleiades Publishing, Inc.