Technical Physics, Vol. 47, No. 1, 2002, pp. 1–5. Translated from Zhurnal Tekhnichesko ĭ Fiziki, Vol. 72, No. 1, 2002, pp. 3–7.

THEORETICAL AND MATHEMATICAL PHYSICS

Transition Radiation from an Extended Bunch of Charged Particles

B. M. Bolotovskiĭ and A. V. Serov

Lebedev Institute of Physics, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia

e-mail: serov@sgi.lpi.msk.ru

Received April 12, 2001

Abstract—The spatial distribution of the transition radiation generated by an extended system of charges is studied. The charges sequentially cross the interface at equispaced points. Also, the transition from a spherical bunch with the uniform charge distribution is considered. The radiation patterns produced by the point charge and by the bunch of the charges are shown to differ significantly at certain sizes of the bunch expressed in terms of the wavelength. Charge distributions such that the transition radiation exhibits the properties typical of the Doppler effect or of Vavilov–Cherenkov radiation are found to be possible. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known that a fast charged particle crossing the interface between two media generates transition radiation [1, 2]. The properties of the transition radiation produced by a single point charge have been studied sufficiently well. When the interface is crossed by an extended bunch of charged particles, the resultant transition radiation may be significantly different from that excited by a single particle. The transition radiation finite-size bunches was considered, e.g., in [2]. As a rule, the authors focus on the conditions under which the radiation generated by a bunch of N particles is the same as that of a point particle with the charge eN. Such studies have formulated certain requirements for the bunch size. Another problem of interest is how to select the bunch sizes for transition radiation to have desired properties (for example, a specified angular or spectral distribution). The maximum coherence may be unachievable in this case in the sense that the radiation intensity from a bunch will not be N^2 times that from a single particle.

As is known, when a charged particle escapes the plane surface of a metal into free space, the field radiated at a certain frequency can be written as

$$E(\omega) = \frac{q}{\pi c R}\left(\frac{\beta\sin\theta}{1-\beta^2\cos^2\theta}\right)\exp\left(i\frac{\omega}{c}R\right), \qquad (1)$$

where q is the charge, $\beta = v/c$ is the ratio between the particle velocity v and the velocity of light c, θ is the angle between the normal to the surface and the vector $\mathbf{R}$ that connects the point where the charge crosses the interface to the point of observation, and $R = |\mathbf{R}|$. The electric field lies in the plane passing through the particle's trajectory and the observation point and is orthogonal to the vector $\mathbf{R}$.

As follows from this formula, when the particle velocity is close to that of light, the radiation is highly directional. It is totally concentrated near a conical surface whose axis is normal to the interface and the vertex angle is on the order of $1/\gamma$, where $\gamma = (1-\beta^2)^{-1/2}$ is the reduced particle energy (the particle energy divided by the rest energy).

However, when the interface is crossed by an extended bunch rather than by one particle, the resultant radiation may significantly differ from that of the particle, as was noted above.

Let the transition radiation be generated by two particles crossing the interface at two points at two different time instants. Then, the radiated field is the superposition of two fields:

$$E_1(\omega) = \frac{q}{\pi c R_1}\left(\frac{\beta\sin\theta_1}{1-\beta^2\cos^2\theta_1}\right)\exp\left(i\frac{\omega}{c}R_1\right), \qquad (2)$$

$$E_2(\omega) = \frac{q}{\pi c R_2}\left(\frac{\beta\sin\theta_2}{1-\beta^2\cos^2\theta_2}\right)\exp\left(i\frac{\omega}{c}R_2 - i\omega t\right), \qquad (3)$$

where t is the time interval between the first and the second bursts of radiation.

Far away from the points of transition, the field radiated can be written as

$$E(\omega) = E_1(\omega) + E_2(\omega) = \frac{q}{\pi c}\left(\frac{\beta\sin\theta}{1-\beta^2\cos^2\theta}\right) \times\left(\frac{1}{R_1}\right)\exp\left(i\frac{\omega}{c}R_1\right)\left[1+\exp\left(i\frac{\omega}{c}\mathbf{nL} - i\omega t\right)\right], \qquad (4)$$

where **L** is the distance between the points at which the particles cross the interface.

In the particular case $\mathbf{nL} - \omega t = 2\pi n$, the field at the point of observation doubles and the energy becomes four times larger than for the case of the single particle. In another particular case, when $\mathbf{nL} - \omega t = \pi(2n + 1)$, the amplitude of the total field vanishes, because the waves generated by the two sources arrive at the point of observation in antiphase and cancel each other. In particular, the field radiated can be reduced to zero in the direction where the radiation generated by a single particle is maximum. In this situation, the radiation emitted at other angles becomes significant; i.e., the radiation pattern changes significantly.

RADIATION GENERATED BY AN INTERMITTENT SOURCE

Consider the case when the interference pattern is formed by a greater number of charged particles and the interference of the transition radiation leads to a pattern typical of the Vavilov–Cherenkov radiation or of the Doppler effect. Let the x axis be directed along a perfectly conducting metal surface (Fig. 1). We assume that an electron strikes the point $x = 0$ at the time instant $t = 0$, producing a burst of transition radiation. Far away from the point $x = 0$, the field can be written as

$$\mathbf{E}_1 = \mathbf{E}(\mathbf{k})e^{i\mathbf{kr} - i\omega t}. \tag{5}$$

Let the second electron, moving with the same velocity, strike the surface at the point $x = L$ at the time instant $t = T$, producing the second burst of transition radiation. Far away from the point $x = L$, the second field can be written as

$$\mathbf{E}_2 = \mathbf{E}(\mathbf{k})e^{i\mathbf{k}(\mathbf{r} - \mathbf{L}) - i\omega(t - T)}. \tag{6}$$

This process is assumed to be periodic; i.e., the nth electron strikes the surface at the point $x_n = x + (n - 1)L$ at the time instant $t_n = t + (n - 1)T$, etc. The total field is

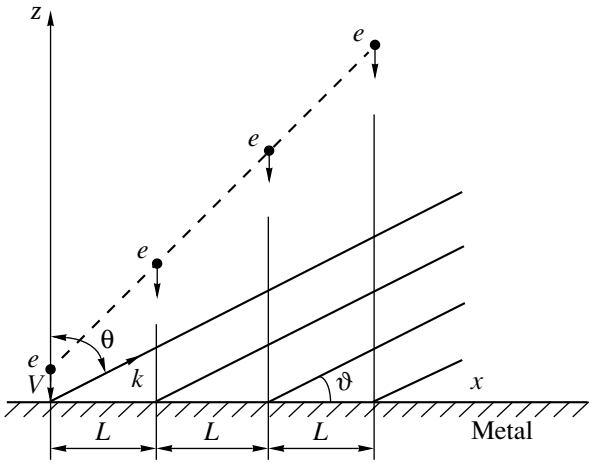


Fig. 1.

the superposition of the fields $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_n$. If the number of the particles is infinite, the sum of these fields is

$$\begin{aligned}\mathbf{E} &= \sum_0^N \mathbf{E}(\mathbf{k}) \\ &= \mathbf{E}(\mathbf{k})e^{i\mathbf{kr} - i\omega t}[1 + e^{i(\mathbf{kL} - \omega T)} + e^{2i(\mathbf{kL} - \omega T)} + \ldots] \\ &= 2\pi\mathbf{E}(\mathbf{k})e^{i\mathbf{kr} - i\omega t}\delta(\mathbf{kL} - \omega T - 2\pi m),\end{aligned} \tag{7}$$

where m is an arbitrary integer.

The total transition radiation field thus differs from the field produced by a single particle by the factor $\delta(\mathbf{kL} - \omega T - 2\pi m)$. In this case, the additional condition is placed on the field radiated: the argument of the delta function should be zero. This means that the transition radiation contains only those waves satisfying the condition $\omega T - \mathbf{kL} = 2\pi m$. Taking into account the geometry of the problem, this condition can be recast as

$$\omega = \frac{2\pi m}{T\left(1 - \dfrac{L}{cT}\cos\vartheta\right)}, \tag{8}$$

where ϑ is the angle between the x axis and the direction of the radiation.

The resulting radiation can be viewed as being produced by an intermittent source that moves along the x axis with the speed $v = L/T$ and flashes with the period T. If $m = 0$ in formula (8), the radiation with the frequency ω can be nonzero only when

$$1 - \frac{L}{cT}\cos\vartheta = 0. \tag{9}$$

In this case, the radiation is other than zero on a cone with a vertex angle ϑ such that $\cos\vartheta = c/v$. This is the condition for the occurrence of the Vavilov–Cherenkov radiation in a vacuum. It should be stressed that, unlike the velocity of bodies, the quantity $v = L/T$ is, in essence, the velocity of a light spot and may exceed the velocity of light. When $m \neq 0$, condition (8) specifies the frequency spectrum typical of the Doppler effect. The formula for the transition radiation field thus contains the terms typical of the Vavilov–Cherenkov radiation and the Doppler effect.

If the number of particles that cross the surface is finite, sum (7) contains a finite number of terms and is reduced to

$$S_N = \frac{1 - e^{iN\Delta}}{1 - e^{i\Delta}} = e^{i\frac{(N-1)\Delta}{2}}\left(\frac{\sin\dfrac{N\Delta}{2}}{\sin\dfrac{\Delta}{2}}\right), \tag{10}$$

where N is the number of particles and $\Delta = \mathbf{kL} - \omega t$.

In this case, the radiation is seen to have a number of maxima instead of a single maximum as in the case of a single particle.

RADIATION PRODUCED BY A SPHERICAL BUNCH

If the transition radiation occurs when the interface is crossed by a bunch of charged particles, its properties depend on the shape of the bunch and the charge distribution in it. In this case, certain information on the radiation can be gained by studying the transition radiation produced by a spherical bunch with the uniform charge distribution. Let a spherical bunch of radius r_0 move along the z axis with a constant speed v and let the initial time instant be that when the bunch touches the interface. We split the bunch into thin layers parallel to the interface. As the bunch moves, these layers successively cross the interface and generate the transition radiation. The total radiation from the bunch is the superposition of the fields produced by each of the layers. The charge on these disk-shaped layers is distributed uniformly. The radius $a(t)$ of the disk varies with time. Initially, the radius equals zero. At a time instant t, the radius is given by

$$a(t) = \sqrt{r_0^2 - (r_0 - vt)^2}. \quad (11)$$

The time variation of the radius defines the evolution of the radiating region. When the spherical bunch touches the interface, the radius of the disk becomes equal to the bunch radius and, subsequently, goes to zero. Consider the rate of charge of the radiating region radius.

Formula (11) yields

$$\frac{da}{dt} = \frac{v(r_0 - vt)}{\sqrt{2r_0vt - v^2t^2}}. \quad (12)$$

From (12), it follows that, initially, the rate at which the radiating region radius grows is higher than the velocity of light. At $t = 0$, the rate da/dt is infinite. Subsequently, as the bunch moves, the rate of expansion of the radiating region decreases, although the region continues expanding until the bunch center crosses the interface. At this time instant, the rate of expansion vanishes; then, the radius starts decreasing and becomes zero at $t = 2r/v$. Immediately before this instant, the radiating region shrinks to the center so that its radius decreases at a rate higher than the velocity of light.

The radiation produced by the spherical bunch crossing the interface can thus be visualized as coming from a certain region of the interface. As the bunch moves, this region appears, expands (at certain moments faster than the velocity of light), then shrinks, and finally disappears. If the speed of the bunch is close to the velocity of light, the radiating region expands with a speed exceeding the velocity of light within the time interval $t \simeq 0$–$3r/c$. Between $t \simeq 1.7r/c$ and $2r/v$, the boundary of the region also moves to the center with a speed exceeding the velocity of light.

Consider the contribution of a disk of radius $a(t)$ to the radiation of the spherical bunch. The geometry of the problem is shown in Fig. 2. The xy plane coincides with the interface. The disk of radius $a(t)$ represents that layer of the spherical bunch crossing the interface at a time instant t. Each of the disk elements is a source of radiation. Consider the field at a point D situated at a large distance R_0 from the disk center ($R_0 \gg a$). The unit vector n pointing from the disk center to the point D makes an angle θ with the normal to the interface (that is, with the z axis). Without loss in generality, the point D can be assumed to lie in the xz plane. Consider a small element $dS = rdrd\varphi$ of the radiating region located at a distance r from the origin and defined by an angle φ. Let R be the distance from the radiating element to the point D. The field of the transition radiation produced by the element dS at the point D can be written as

$$E(\omega) = \frac{\sigma}{\pi cR}\left(\frac{\beta\sin\theta}{1-\beta^2\cos^2\theta}\right)\exp\left(i\frac{\omega}{c}R\right)dS, \quad (13)$$

where σ is the surface charge density.

The total field is determined by integrating over the disk of radius a:

$$E(\omega) = \frac{\sigma}{\pi c}\int_0^a rdr\int_0^{2\pi}\left(\frac{\beta\sin\theta}{1-\beta^2\cos^2\theta}\right)\frac{1}{R}\exp\left(i\frac{\omega}{c}R\right)d\varphi. \quad (14)$$

Since $R_0 \gg a$, we have

$$R = R_0 - \mathbf{nr} = R_0 - r\sin\theta\sin\varphi.$$

With this relationship, the integration gives the expression

$$E(\omega) = \frac{\sigma}{\pi}\left(\frac{\beta\sin\theta}{1-\beta^2\cos^2\theta}\right)\frac{a(t)J_1\left(\frac{\omega}{c}\sin\theta a(t)\right)}{\omega\sin\theta}. \quad (15)$$

The angular field distribution is described by two factors. The factor in the parentheses describes the tran-

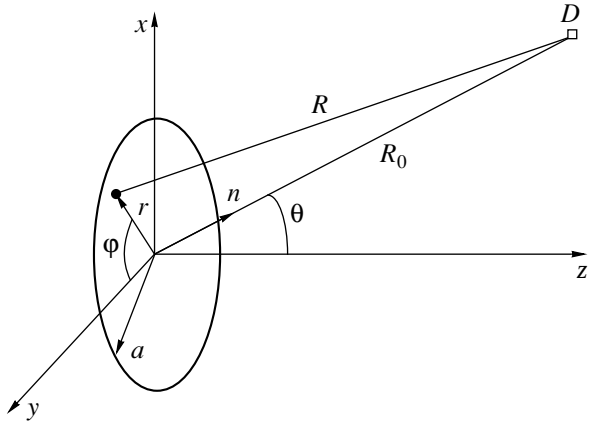


Fig. 2.

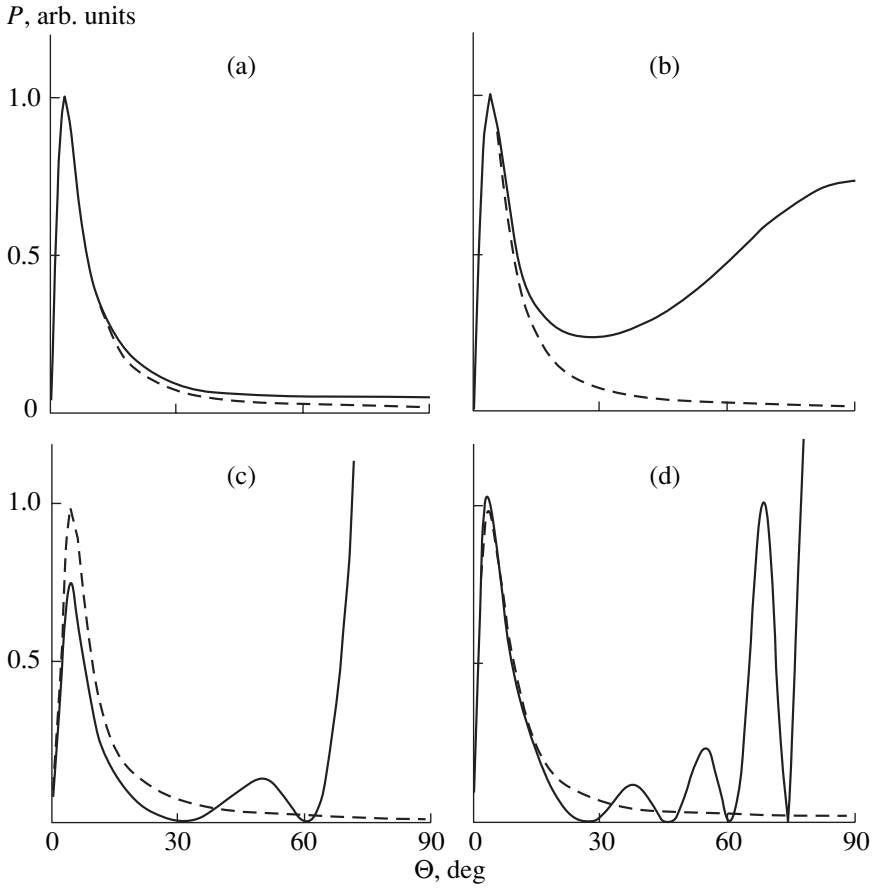


Fig. 3. Transition radiation patterns at λ/r_0 = (a) 1.5, (b) 0.85, (c) 0.35, and (d) 0.2.

sition radiation of a unit point charge; the second factor allows for the interference of the fields produced by different parts of the circular region.

As the spherical bunch passes through the interface, the radius of the radiating region changes, so that the total field is given by the integral over the time it takes for the bunch to cross the interface (or, which is the same, by the integral over all the layers). Ultimately,

$$E(\omega) = \frac{\sigma}{\pi}\left(\frac{\beta\sin\theta}{1-\beta^2\cos^2\theta}\right)\int_0^{2r/v}\sqrt{2r_0vt-v^2t^2}\times\frac{J_1\left(\frac{\omega}{c}\sin\theta\sqrt{2r_0vt-v^2t^2}\right)}{\omega\sin\theta}\exp\left(i\frac{\omega}{c}t\right)dt. \quad (16)$$

It can be easily shown that the integral is the Fourier transform of the charge density distribution over the spherical bunch.

Figure 3 shows the transition radiation intensity $P(\omega) \sim E^2(\omega)$ versus angle θ calculated from formula (16) at different ratios of the wavelength λ to the sphere radius r_0. The curves are normalized to the maximum intensity. Note that the sphere is a simple body characterized by one size. Figure 3a refers to the case when the wavelength exceeds the sphere radius. The solid line depicts the field produced by the sphere; the dashed line, by the point source. In this case, the radiation pattern of the bunch is seen to differ little from that of the point source, except at large angles. Our calculations show that the difference decreases with increasing ratio λ/r_0. For λ/r_0 = 0.85 (Fig. 3b), the radiation of the extended bunch significantly differs from that of the point charge at large angles. With decreasing λ/r_0, the small-angle radiation is suppressed, while the large-angle radiation becomes significant. Figures 3c and 3d show the radiation patterns at λ/r_0 = 0.35 and 0.2, respectively, which are seen to contain extra lobes, along with the narrow lobe $\theta \simeq 1/\gamma$. As follows from our calculations, a further decrease in the ratio λ/r_0 raises the number of lobes.

The decrease in the forward radiation may be explained by the fact that the extended bunch contains volume elements that radiate in antiphase. This decrease is particularly noticeable when the bunch length is comparable to or greater than the wavelength.

The radiation enhancement at large angles may be attributed to the coherent directional emission generated by the radiating region expanding at a supraluminal speed. This radiation can be regarded as an analog of the Vavilov–Cherenkov radiation.

We considered charged bunches of the simplest shape. Real bunches of relativistic particles obtained in accelerators have more complex density distributions. Therefore, the angular distribution of the real transition radiation will be different from that considered above. However, it will also differ from the transition radiation produced by a point charge. It should also be noted that the size of the bunch and the particle distribution in it can be judged from the radiation pattern.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 99-02-18183 and 00-02-17381).

REFERENCES

1. V. L. Ginzburg and I. M. Frank, Zh. Éksp. Teor. Fiz. **16**, 15 (1946).
2. V. L. Ginzburg and V. N. Tsytovich, *Transition Radiation and Transition Scattering* (Nauka, Moscow, 1984).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 102–106. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 105–109.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

The Study of Polystyrene–Fullerene Solid-Phase Composites

S. V. Gladchenko, G. A. Polotskaya, A. V. Gribanov, and V. N. Zgonnik

Institute of High-Molecular Compounds, Russian Academy of Sciences, St. Petersburg, 199004 Russia

Received April 27, 2001

Abstract—Polystyrene-based films containing from 0 to 45 mol % of C_{60} fullerene are investigated. It is found that the introduction of fullerene increases the molecular packing density of polystyrene chains, thus affecting the transport of small molecules through the polymer films. Gas diffusion through the composite films is slower than through the polystyrene films, while the gas-distributing properties of the composites are higher. The dielectric relaxation method shows that the solution-grown films retain the cluster state of fullerene. When the films containing more than 0.15 mol % of fullerene are heated above the glass transition temperature of polystyrene away from air, the relaxation time of the α transition in polystyrene is found to increase. This effect is associated with strong chain interaction via fullerene molecules incorporated into polystyrene–C_{60} complexes. The NMR spectroscopy method allows us to observe the typical changes in the composites. © *2002 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

The introduction of chemically active purely carbon fullerene molecules into conventional polymers allows researchers to synthesize materials with new intriguing properties [1–4]. Fullerene-containing polymers under investigation include systems where the polymer and fullerene are linked by one or several covalent bonds, as well as composites where bonding between the polymer matrix and fullerene is due to donor–acceptor and/or hydrophilic–hydrophobic interactions.

Polystyrene (PS) is one of the polymers that have been successfully modified by fullerene [5–20]. Fullerene-containing PS has been produced by radical or anionic polymerization and also from appropriate compositions. Conventionally, fullerene-containing PS is made by the free-radical polymerization of styrene with fullerene present either in the block or in an aromatic solvent. The resulting polymer is colored dark brown [5, 6]. Another way of producing fullerene-containing PS that has been extensively discussed in the literature is anionic polymerization involving the formation of active fullerene-containing centers. With this method, complex stellate and dumbbell-like structures have been produced [7–12]. It has been shown that fullerene and PS may produce compounds with coordination bonds [13, 14]. Works concerned with fullerene-containing PS have been concentrated largely on the synthesis, as well as on the optical (in particular, spectral), physical, and other properties of fullerene-containing PS in the growth solutions [15–18].

However, to estimate the actual potentialities of fullerene-containing PS, one must know its behavior not only in the solutions but also in the solid state and find basic characteristics that govern the latter. The effect of fullerene on the thermal destruction of fullerene–PS composites has been studied by thermal mass-spectrometric analysis [19, 20]. The interaction between the components, as well as the interference of their structural and kinetic parameters, specify the physical properties of the polymer.

The aim of this article is to study the structure of PS–C_{60} composites in the solid state and illuminate their properties of practical significance by using the NMR method, gas diffusion, and the method of dielectric relaxation. The subject of investigation is films obtained from dilute solutions of PS–C_{60} mixtures.

EXPERIMENT

A fullerene extract containing more than 98% C_{60} (Institute of Physics and Technology, Russian Academy of Sciences) was vacuum-heated up to 100°C to remove moisture and dissolved in toluene. PS with a molecular mass of 3.2×10^5 was also dissolved in toluene. Then, both solutions were mixed with a desired proportion [mol % C_{60} = (number of C_{60} molecules) : (number of PS monomer links)]. The films ~30 μm thick were deposited on the cellophane surface by evaporating the toluene from the C_{60} + PS solution at 40°C.

The density ρ of the films was determined by flotation in an aqueous solution of saccharose at 20°C. The molecular packing density coefficient was calculated by the Kitaigorodsky formula

$$k = \frac{N\sum_i \Delta V_i}{M}\rho,$$

1063-7842/02/4701-0102$22.00

where N is the Avogadro number, M is the monomer link mass, and $\sum \Delta V_i$ is the monomer link volume determined from the sum of group contributions [21].

Prior to measuring their insulating properties, the films were dried in a vacuum at 40°C for 4 h. The measurements were performed at frequencies from 0.1 to 100 kHz in an evacuated sealed glass cell with clamping chromium-plated brass electrodes at a pressure of 1.3 Pa and temperatures from 20°C to 130–150°C. After the first measurement, the samples were cooled and the loss tangent $\tan\delta$ was measured again in the same temperature and frequency ranges.

The energy of activation ΔU of dipole polarization was calculated from the Arrhenius relationship $\tau = \tau_0 e^{\Delta U/RT}$. The relaxation times were found from the relationship $\tau = 1/2\pi f_m$, where τ is the relaxation time at which $\tan\delta$ goes through a maximum at a given temperature and f_m is the frequency at which $\tan\delta$ is maximal at a constant temperature.

High-resolution NMR spectra from solid ^{13}C were taken with an SKhR-100 spectrometer with an operating frequency of 25.18 MHz using magic-angle rotation and polarization transfer techniques.

The gas-distributing properties (gas permeation selectivity) were studied with a PGD-01 setup by passing air through the film (membrane). The film was placed in a diffusion cell at a temperature of 30°C and a partial pressure drop of 101 325 Pa. The amount and the composition of the gas at the outlet were analyzed chromatographically. Using the chromatographic data, we calculated the oxygen and nitrogen permeabilities, as well as the separability of the gases offered by this membrane. The permeability of gas i ($\bar{P}_i$) depends on the physicochemical properties of the membrane polymer and is given by

$$\bar{P}_i = \frac{Vl}{A\tau\Delta p},$$

where V is the gas volume, l is the membrane thickness, A is the membrane surface area, τ is the gas penetration time, and Δp is the pressure drop across the membrane.

Selectivity of permeation results from the fact that different gases permeate a membrane at different rates. The separability of two gases is defined by their permeability ratio

$$\alpha_{i/j} = \frac{P_i}{P_j}.$$

EXPERIMENTAL RESULTS

In studying the solutions by scattering of light, it has been found that fullerene and PS produce aggregates with coordination bonds [13, 14]. Usually, the aggregation proceeds either by compacting or, conversely, by loosening the polymer chain configuration. To reveal the structure modifications upon adding C_{60} to PS, we compared the density of the initial PS and that of the PS–0.3 mol % C_{60} composite. From the density values measured (1051 and 1069 kg/m^3, respectively), we calculated the molecular packing density coefficients for the polymer chains, which were found to be 0.662 and 0.668, respectively. Thus, the introduction of C_{60} makes the PS chain packing more compact. A method fairly sensitive to a change in the polymer packing density is the transport of small molecules through polymer films.

Figure 1 shows the measured oxygen and nitrogen permeabilities for pure PS and for PS containing C_{60} in amounts from 0.075 to 0.30 mol %. It is seen that the composites have the lower permeabilities and that the decrease in the permeabilities is accompanied by the increase in the separability of the gases entering from air. Such a modification in the PS transport properties is favorable, since the gas permeation selectivity of unmodified PS needs improvement.

The distribution of C_{60} molecules as a function of the concentration was constructed using the molecular mobility values measured by the dielectric method. The method is based on the fact that the molecular mobility (or the relaxation time of internal polymer parameters) is sensitive to intra- and intermolecular interactions. In our case, intermolecular interactions are of interest, since they are responsible for the difference in the mobilities of macrochains when they are surrounded by the same molecules or by molecules of the other component.

The dielectric properties of pure PS have been well studied [22, 23]. PS is a weakly polar polymer that is in the glassy state at room temperature. The relaxation

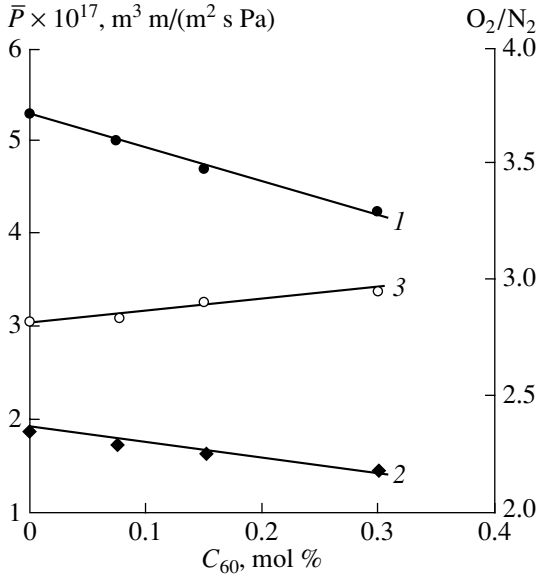


Fig. 1. (*1*) Oxygen and (*2*) nitrogen permeabilities and (*3*) gas separability (O_2/N_2) vs. fullerene content in the PS–C_{60} composites.

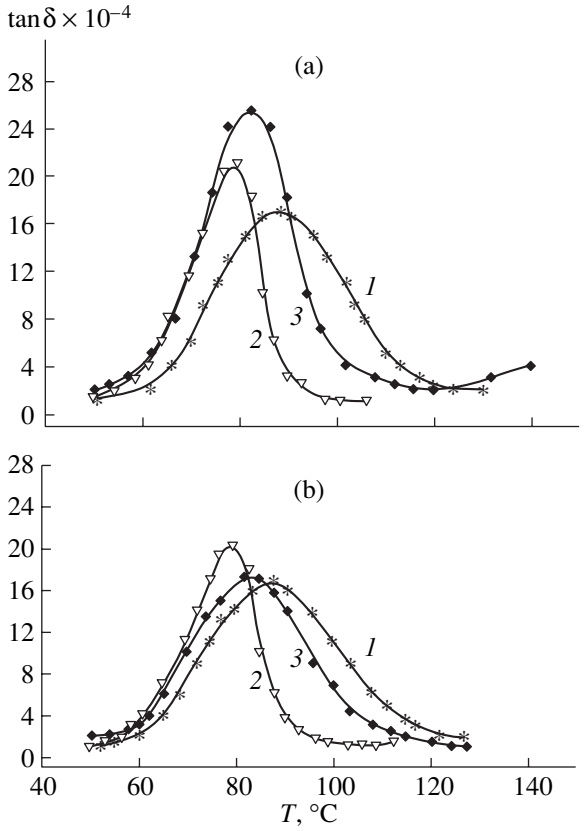


Fig. 2. Temperature dependence of $\tan\delta$ for the composites with the fullerene content (*1*) 0, (*2*) 0.075, and (*3*) 0.15 mol %. (a) As-prepared samples and (b) samples heated to 150°C away from air. The frequency is 0.1 kHz.

dielectric losses are observed only near the rubber transition (α transition). Far away from the absorption peak for this relaxation process, $\tan\delta$ is of the order of 10^{-4}.

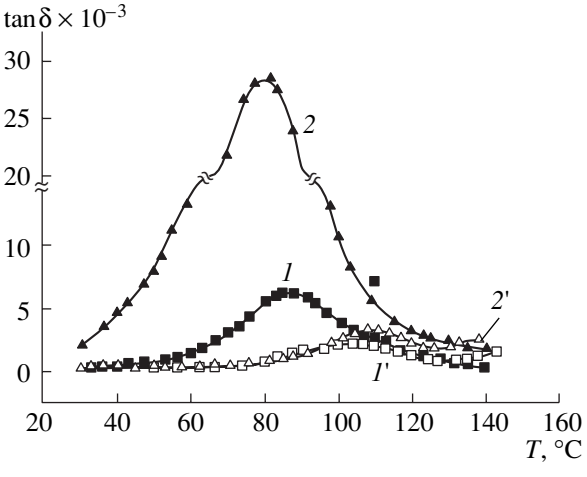


Fig. 3. Temperature dependence of $\tan\delta$ for the composites with the fullerene content (*1*, *1*') 0.30 and (*2*, *2*') 0.45 mol %. (*1*, *2*) As-prepared samples and (*1*', *2*') samples heated to 150°C away from air. The frequency is 0.1 kHz.

Figure 2 presents the temperature dependences of $\tan\delta$ for the pure PS and for the PS containing fullerene in amounts of 0.075 and 0.15 mol %. Each of the tree samples has the single dipole polarization region that is associated with the α transition in the PS. The addition of fullerene somewhat decreases the temperature of the dielectric absorption peak while increasing its intensity. When the samples are heated to temperatures above the PS glass-transition temperature away from air, the temperature position of $\tan\delta_{max}$ remains the same but its peak intensity noticeably changes (Fig. 2b).

For the composites with 0.30 and 0.45 mol % of fullerene, $\tan\delta$ varies with temperature in a much different way (Fig. 3). At temperatures near the PS glass-transition temperature (≈80°C), a new relaxation range of dielectric losses appears. This peak grows with C_{60} content and is much higher than that associated with the α transition in the pure PS. Heating to 130°C suppresses this peak of the dielectric absorption. For the heated PS–0.30 mol % C_{60} and PS–0.45 mol % C_{60} samples, only one peak of $\tan\delta$ is observed. Its height is comparable to that associated with the α transition, but the relaxation times are larger (Fig. 3). For the samples heated, the apparent energy of activation ΔU of the α transition is 65–79 kcal/mol, as follows from the $\log f_{max} = \varphi(1/T)$ dependences (Fig. 4).

Using the data for the samples heated (Figs. 2–4), we constructed the dependences of the energy of activation ΔU, relaxation time τ, and temperature position T_M of $\tan\delta_{max}$ on the fullerene content at f = 0.1 kHz (Fig. 5). The concentration dependences of $\log\tau$ and T_M are nonmonotonic. The relaxation time of the α transition in the pure PS is 10^{-4} s and is reduced to 10^{-5} s upon adding 0.075 mol % C_{60}. As the C_{60} content

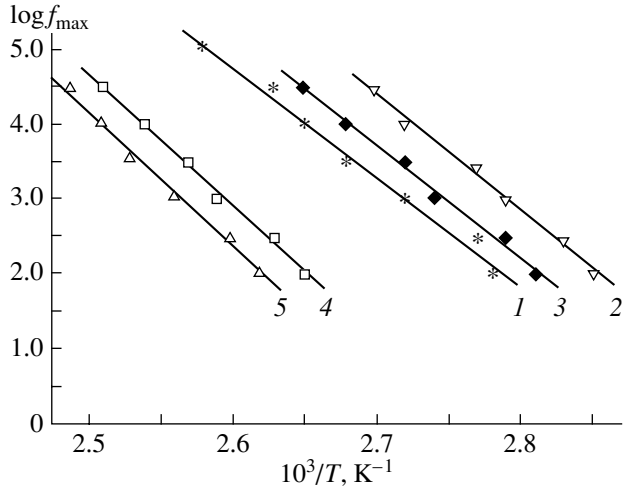


Fig. 4. $\log f_{max} = \varphi(1/T)$ for the composites with the fullerene content (*1*) 0, (*2*) 0.075, (*3*) 0.15, (*4*) 0.30, and (*5*) 0.45 mol % heated to 150°C.

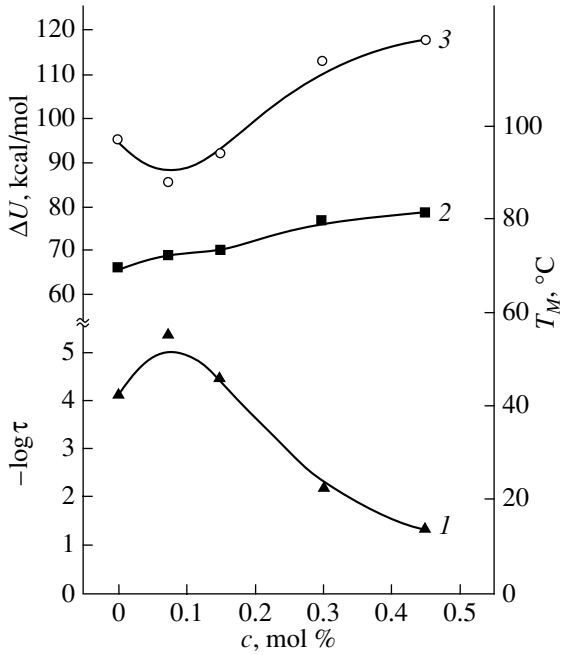


Fig. 5. (*1*) Relaxation time, (*2*) energy of activation, and (*3*) temperature position of $\tan\delta_{max}$, T_M, vs. fullerene content in the composite after heating to 150°C. The frequency is 0.1 kHz.

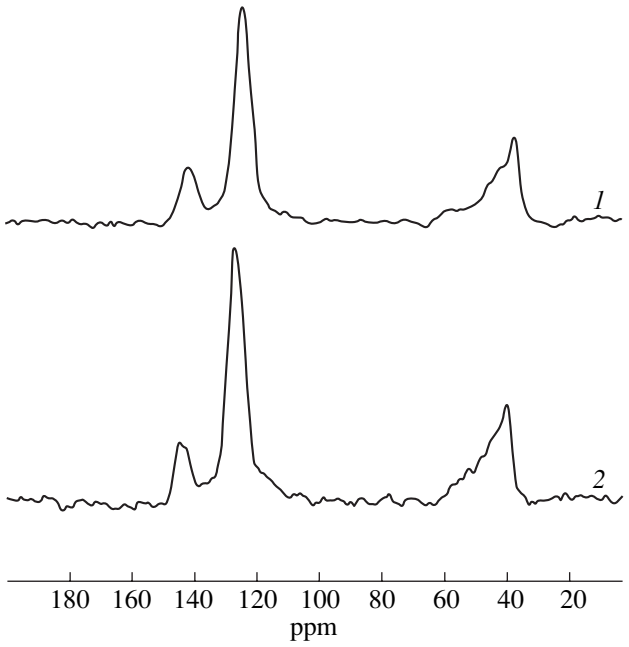


Fig. 6. NMR spectra from the (*1*) as-prepared and (*2*) heated (100°C) PS–0.45 mol % C_{60} samples.

increases further, so does τ. For the sample with 0.45 mol % C_{60}, the relaxation time of the α transition becomes equal to 5×10^{-2} s (T = 100°C). The energy of activation changes by ≈14 kcal/mol and tends to increase.

The typical modifications of the PS–C_{60} composites during thermal treatment are also indicated by NMR spectra. The absorption band of aromatically substituted carbon in a PS molecule is known to be close to that of C_{60} (≈145 ppm). In the PS–0.075 mol % C_{60} composite, the signal of C_{60} is practically indistinguishable against the background of the intense signal from this carbon in the PS. However, when the C_{60} content grows to 0.45 mol %, the presence of the C_{60} molecules is indicated by the minor distortion of the signal on the strong-field side at 140–150 ppm (Fig. 6, curve *1*). This seems to be quite reasonable because the amount of the fullerene is relatively small and also because it becomes to a great extent immobile when its molecules are uniformly distributed among the PS macromolecules. After the sample has been heated to 100°C, the signal at 145 ppm splits up, the exact position of the peak corresponding to the fullerene signal, 144.3 ppm (Fig. 6, curve *2*). It is likely that after heating to 100°C, specific conditions are set in the PS–C_{60} system under which the rotational mobility of the C_{60} molecules increases.

DISCUSSION

We assume that the above data reflect PS–fullerene molecular interactions. In the presence of fullerene, the PS molecular packing density increases, which affects the transport of small molecules through polymer films. The gas molecules diffuse through the composite films more slowly than through the PS films and the permeation selectivity of the former is improved.

The relaxation properties of the PS–C_{60} composites were studied with the dielectric method by comparing the $\tan\delta = \varphi(T)$ dependences for the as-prepared and heated samples. It is known [24] that fullerenes may form clusters when in the solution. We assume that fullerene clusters are also present in the solid-phase PS–C_{60} composites before heating. If so, the discrepancy between the $\tan\delta = \varphi(T)$ dependences for the pure PS and the composites is related to the dipole mobility in fullerene clusters, whose kinetic characteristics are specified largely by fullerene–fullerene interaction. The mechanism of dipole relaxation in fullerenes still remains unclear. It has been argued [25, 26] that the oxygen impurity is to a great extent responsible for this process. On the other hand, other researchers insist that the dielectric properties of fullerenes depend on impurities (external factors) insignificantly [27, 28].

Upon heating, fullerene clusters are likely to decompose, of which the absence of the intense relaxation transition in the $\tan\delta = \varphi(T)$ curve for the samples heated is an indication. Also, it appears that the decomposition of fullerene clusters at temperatures when PS is in the rubbery state favors the homogeneous mixing

of the components, thus causing PS–C_{60} aggregates to occur. Because of this, in the composites containing 0.075 and 0.15 mol % of C_{60}, the fullerene-assisted plasticization of the polymer takes place (Figs. 2 and 4). The C_{60} molecules embed in the interchain space, loosening the chain links and reducing the PS glass-transition temperature. Finally, the increase in the relaxation time of the α transition with growing C_{60} content can be explained as follows. C_{60} molecules can strongly interact with each other because of a large density of conjugated bonds [29]. As the C_{60} content in PS grows, the possibility of interaction between C_{60} molecules involved in PS–C_{60} aggregates increases, which may lead to cross linking in PS.

REFERENCES

1. H. W. Kroto, A. W. Allaf, and S. P. Balm, Chem. Rev. **91**, 1213 (1991).
2. A. Hirsch, *The Chemistry of the Fullerenes* (G. Thieme, Stuttgart, 1994).
3. E. N. Karaulova and E. I. Bagriĭ, Usp. Khim. **68**, 979 (1999).
4. R. F. Curl, Usp. Fiz. Nauk **168**, 331 (1998).
5. H. Okamura, N. Ide, M. Minoda, *et al.*, Macromolecules **31**, 1859 (1998).
6. D. Stewart and C. T. Imrie, J. Chem. Soc., 1383 (1996).
7. V. N. Zgonnik, E. Yu. Melenevskaya, L. S. Litvinova, *et al.*, Vysokomol. Soedin., Ser. A **38**, 203 (1996).
8. E. Yu. Melenevskaya, L. V. Vinogradova, L. S. Litvinova, *et al.*, Vysokomol. Soedin., Ser. A **40**, 247 (1998).
9. T. Cao and S. E. Webber, Macromolecules **28**, 3741 (1995).
10. Y. Ederle and C. Mathis, Macromolecules **32**, 2546 (1999).
11. Y. Ederle and C. Mathis, Macromolecules **32**, 554 (1999).
12. E. Cloutet, J.-L. Fillaut, D. Astrue, and Y. Gnanou, Macromolecules **32**, 1043 (1999).
13. M. L. Sushko, L. V. Vinogradova, E. E. Kever, *et al.*, in *Abstracts of 3rd International Symposium "Molecular Order and Mobility in Polymer Systems," St. Petersburg, 1999*, p. 105.
14. M. L. Sushlo, V. O. Aseyev, H. Tenhu, and S. I. Klenin, Mol. Mater. **13**, 339 (2000).
15. C. Weis, C. Friedrich, R. Mulhaupt, and H. Frey, Macromolecules **28**, 403 (1995).
16. Y.-P. Sun, B. Ma, C. E. Bunker, and B. Liu, J. Am. Chem. Soc. **117**, 12 705 (1995).
17. D. A. Loy and A. A. Roger, J. Am. Chem. Soc. **114**, 3977 (1992).
18. A. G. Camp, A. Lary, and W. T. Ford, Macromolecules **28**, 7959 (1995).
19. V. Zgonnik, L. Shibaev, T. Antonova, *et al.*, in *Proceedings of the 3rd International Workshop on Fullerenes and Atomic Clusters, St. Petersburg, 1997*, p. 4.
20. B. M. Ginzburg, A. O. Pozdnyakov, O. F. Pozdnyakov, and B. P. Redkov, Zh. Prikl. Khim. (St. Petersburg) **73**, 484 (2000).
21. A. A. Askadskiĭ, *Structure and Properties of Heat-Resistant Polymers* (Khimiya, Moscow, 1981).
22. O. Yano and Y. Wada, J. Polym. Sci., Part A-2 **9**, 669 (1971).
23. B. N. Satin, T. P. Orlova, and L. N. Lobanov, Vysokomol. Soedin., Ser. A **10**, 1921 (1968).
24. V. N. Bezmel'nitsin, A. V. Eletskiĭ, and M. V. Okun', Usp. Fiz. Nauk **168**, 1195 (1998) [Phys. Usp. **41**, 1091 (1998)].
25. M. S. Dresselhaus, G. Dresselhaus, and P. S. Eklund, *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1996).
26. A. Pevzner, A. F. Hebard, and M. S. Dresselhaus, Phys. Rev. B **55**, 16 439 (1997).
27. A. A. Pronin, S. V. Demishev, N. E. Sluchanko, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 573 (2000) [Phys. Solid State **42**, 586 (2000)].
28. P. Mondal, P. Lunkenheimer, and A. Loidl, Z. Phys. B **99**, 527 (1996).
29. L. A. Gribov and B. K. Novosadov, Zh. Strukt. Khim. **40**, 582 (1999).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 107–113. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 110–116.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

The Minimization of Residual Stresses in Magnetic Systems with Quasi-Force-Free Current Distribution

G. A. Shneerson, O. S. Koltunov, and V. Yu. Khozikov

St. Petersburg State Technical University, ul. Politekhnicheskaya 29, St. Petersburg, 195251 Russia

e-mail: integr@delfa.net

Received June 14, 2001

Abstract—The feasibility of a magnet experiencing very low mechanical stresses is studied. Megagauss fields can be generated without failure of the solenoid winding and excessively increasing the winding dimensions if the current distribution is force-free in its inner part and the axial current closes in its outer part. The approximation of the quasi-force-free current distribution by discrete conductive layers is considered. Analysis and computations are performed for two force-free winding approximations. In the former case, the conductors are divided into N paired balanced layers where the azimuth and poloidal currents, respectively, pass. The equal and oppositely directed forces arising in the layers of a pair are transferred to insulating spacers between the layers. Stresses in the spacers can be reduced down to values N times as low as the magnetic field pressure $B_0^2/(2\mu_0)$ at the solenoid axis. If the current direction in each of the layers meets the balance condition, the residual stresses can be N^2 times lower than $B_0^2/(2\mu_0)$, because the tensile and compressive forces in the layers partially cancel. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Mechanical stresses in magnetic systems with discrete current layers can significantly be reduced if the current distribution over the layers is almost force-free (so-called force-free systems). The problem of minimizing residual electromagnetic forces is of great importance both in the production of magnets generating superhigh magnetic field and for increasing the superconductor critical current.

An axisymmetric quasi-force-free solenoid of a finite length (Fig. 1) can be implemented as a system of layers where the azimuth and poloidal currents pass. In the continuous current distribution approximation (i.e., with the infinite number of layers), the field in region *1* should be force-free. On the outside, this region borders region *2*, where the axial current δ_z closes. In this region, the field is not force-free, their induction having only the azimuth component.

It has been shown [1–4] that, with the axial current distribution in region *2* appropriately selected, the outer radius of the magnetic system, R_2, and that of the winding, R_1, relate as $R_2/R_1 \approx (2B_0^2/B_M^2 + 1)^{1/2}$, where B_0 is the induction at the axis and $B_M = \sqrt{2\mu_0\sigma_0}$ is the so-called magnetic ultimate strength, which depends on the mechanical ultimate strength σ_0 of the material. Note in pass that in standard solenoids, having a uniformly loaded winding with the azimuth current [5, 6], the outer-to-inner radius ratio grows as $\exp(B_0^2/B_M^2)$, so that the winding dimensions become intolerably large if $B_0/B_M \gg 1$.

Let us try to approximate the force-free distribution by a set of conductors in the region far away from the end faces, where the radial electromagnetic forces are the highest. In this area, the radial components of the

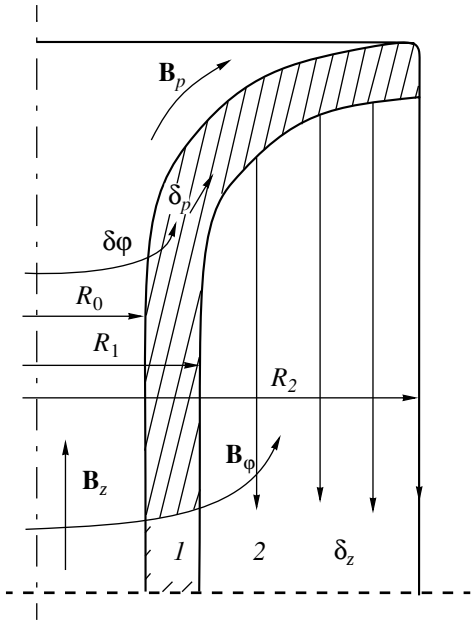


Fig. 1. Finite-length solenoid with quasi-force-free winding *1* and region *2* of oppositely directed current.

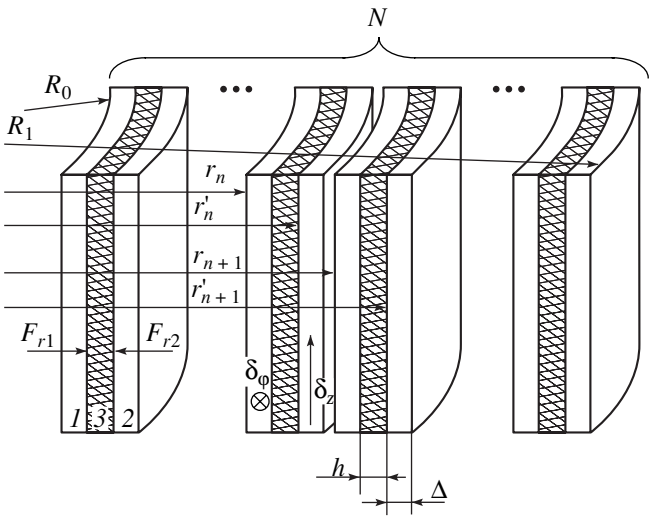


Fig. 2. Quasi-force-free winding with paired balanced current layers.

current density and of the induction are absent. The problem thus stated (the edge fields are ignored) is of independent significance for a superconducting cable [7].

The design of the end face of a quasi-force-free magnet calls for special investigation (see, e.g., [4], where the design of a thin winding is considered).

The possibility of sharply decreasing the volume electromagnetic forces has been demonstrated in [2, 7] by the example of idealized systems with discrete current layers of a vanishingly small thickness. In this work, we approximate the force-free distribution by a set of conductors of a small yet finite thickness.

QUASI-FORCE-FREE WINDING WITH PAIRED BALANCED CURRENT LAYERS (THE NUMBER OF PAIRS $N \gg 1$)

Let the winding consist of N pairs of layers with orthogonal current in each of the pairs (Fig. 2). The layers are separated by insulating spacers. The distribution of the axial and azimuth currents that approximates the force-free current distribution is such that, in each of the pairs, the tensile force acting on inner turn *1* with the azimuth current is counterbalanced by the compressive force acting on outer turn *2* with the axial current. The pairs are balanced, and insulating spacer *3* is under compression.

We assume that the thicknesses of layers *1* and *2* in all the pairs are the same. Then, for the nth pair, the radial force per unit surface area of layer *1* with the azimuth current is given by

$$F_{r1} = (B_z^2(r_n) - B_z^2(r_{n+1}))/(2\mu_0), \qquad (1)$$

where B_z is the axial component of the induction, for which we have

$$B_z(r_{n+1}) = B_z(r_n) - \mu_0 \int_{r_n}^{r_n+\Delta} \delta_\varphi dr \approx B_z(r_n) - \mu_0 \delta_\varphi (\Delta - \Delta^2/(2r_n)). \qquad (2)$$

Equation (2) implies that $\delta_{\varphi,n} = \delta_\varphi(r_n)$, where δ_φ is given by

$$\delta_\varphi = \delta_{\varphi,n} r_n / r. \qquad (3)$$

The terms of the order of $\mu_0 \delta_\varphi \Delta^3 / r_n^2$ are rejected. Without these terms, the force F_{r1} is expressed as

$$F_{r1} \approx B_z(r_n)\delta_{\varphi,n}\Delta - B_z(r_n)(\delta_{\varphi,n}\Delta^2/(2r_n)) - \mu_0 \delta_{\varphi,n}^2 \Delta^2/2. \qquad (4)$$

The force acting on layer *2* is defined by the formula

$$F_{r2} = -\int_{r'_n}^{r'_n+\Delta} \delta_{z,n} B_\varphi dr,$$

where

$$r'_n = r_n + h_1 + \Delta,$$

$$B_\varphi(r) = \frac{r'_n}{r} B_\varphi(r'_n) + \frac{\mu_0 \delta_{z,n}}{2}\left(\frac{r^2 - r_n'^2}{r}\right).$$

Hence,

$$F_{r2} \approx -\left[B_\varphi(r'_n)\delta_{z,n}\Delta - B_\varphi(r'_n)\delta_{z,n}\frac{\Delta^2}{2r'_n} + \mu_0 \delta_{z,n}^2 \frac{\Delta^2}{2}\right], \qquad (5)$$

where the terms of the order of $B_\varphi(r'_n)\delta_{z,n}\Delta^3/r_n'^2$ and $\mu_0 \delta_{z,n}^2 \Delta^3 / r'_n$ are rejected.

The balance condition $F_{r1} + F_{r2} = 0$ should be met in each of the pairs. Eventually, we arrive at the set of equations that allows us to find the distribution of the azimuth currents over the layers for a given distribution of the axial currents or to solve the inverse problem:

$$B_z(r_n)\delta_{\varphi,n} - B_\varphi(r'_n)\delta_{z,n} - \left(\frac{B_z(r_n)\delta_{\varphi,n}}{r_n} - \frac{B_\varphi(r'_n)\delta_{z,n}}{r'_n}\right)\frac{\Delta}{2} - \mu_0 \Delta(\delta_{\varphi,n}^2 + \delta_{z,n}^2)/2 = 0. \qquad (6)$$

In these equations, the induction components B_φ and B_z are determined from the following approximate formulas:

$$B_z(r_n) \approx B_0 - \Delta\mu_0 \sum_1^{n-1} \delta_{\varphi,n}\left(1 - \frac{\Delta}{2r_n}\right), \qquad (7)$$

$$B_\varphi(r'_n) \approx \frac{r'_{n-1}}{r'_n}\left[B_\varphi(r'_{n-1}) + \mu_0\delta_{z,n-1}\Delta\left(1 + \frac{\Delta}{2r'_{n-1}}\right)\right]. \quad (8)$$

The absolute values of the radial and azimuth stresses in spacer *3* are calculated from the expression

$$\begin{aligned}|\sigma_\varphi| &= |\sigma_z| = |F_{r1}| = |F_{r2}| \\ &\approx B_\varphi(r'_n)\delta_{z,n}\Delta(1 - \Delta/(2r'_n)) + \mu_0\Delta^2\delta^2_{z,n}/2.\end{aligned} \quad (9)$$

In (9), the first term is principal. It has the order of $B_0^2/(\mu_0 N)$, since B_φ and B_0 are of the same order and δ_φ has the order of $2B_0/(\mu_0 d)$, where $d = 2N\Delta$ is the thickness of the winding. Thus, the compressive stresses in the orthogonal-current system is roughly N times as low as the magnetic pressure corresponding to the field B_0. The actual values of the compressive forces in various layers can be obtained by solving the set of equations (7) and (8). This set is rather awkward if $N \gg 1$; therefore, the terms of the order of Δ/r and $1/N$ should be neglected when optimization problems are solved. Then, only the first two terms are left on the left of Eq. (6). Next, it is appropriate to approximate the actual induction distribution (broken line) by a continuous differentiable curve $B_{\varphi,z}(r)$ passing through the values of $B_{\varphi,z}(r_n)$. With such an approach, we have the approximate expressions

$$\begin{aligned}B_z(r_{n+1}) &\approx B_z(r_n) - \mu_0\delta_{\varphi,n}\Delta \approx B_z(r_n) \\ &\quad + (dB_z/dr)2(\Delta + h), \\ r'_{n+1}B_\varphi(r'_{n+1}) &\approx r'_nB_\varphi(r'_n) + 2\mu_0 r'_n\delta_{z,n}\Delta \\ &\approx r'_nB_\varphi(r'_n) + 2(\Delta + h)d(rB_\varphi)/dr,\end{aligned}$$

which involve the derivatives of the continuous functions $B_z(r)$ and $B_\varphi(r)$.

From these expressions, we obtain those for the current density at the edges of the layers:

$$\begin{aligned}\delta_{\varphi,n} &\approx -\frac{2(\Delta + h)}{\mu_0\Delta}\left(\frac{dB_z}{dr}\right)_{r=r_n}, \\ \delta_{z,n} &\approx \frac{2(\Delta + h)}{\mu_0\Delta}\frac{1}{r}\left[\frac{d(rB_\varphi)}{dr}\right]_{r=r_n}.\end{aligned} \quad (10)$$

Eventually, instead of (6), we come to the equation that relates the axial and azimuth components of the induction of the field that approximates the quasi-force-free field:

$$B_\varphi\frac{1}{r}\frac{d}{dr}(rB_\varphi) + B_z\frac{dB_z}{dr} = 0. \quad (11)$$

As follows from (11), this field is strictly force-free. Given the distribution of one of the induction components, we can find the other with (11). Such a computing procedure is suitable for finding a current distribution that should meet some specific requirements. Once the "smoothed" induction or current distribution has been derived, one can estimate the stresses in a real (discrete) set of conductors. Although the exact current and voltage values in the layers can be obtained by numerically solving set (6)–(8), the approximate solutions of the smoothed equations are of interest, because sometimes they are derived analytically.

For example, let us find the current distribution such that all of the layers are equiloaded:

$$|\sigma_\varphi| = |\sigma_r| = \sigma_0 = B_\varphi\delta_{z,n}\Delta. \quad (12)$$

From (12), we obtain the equation for B_φ:

$$\frac{\sigma_0(\Delta + h)}{2(\Delta + h)} = \frac{B_\varphi}{\mu_0}\frac{1}{r}\frac{d}{dr}(rB_\varphi). \quad (13)$$

Then, we assume for simplicity that $h \ll \Delta$. Equation (12) allows us to find the azimuth component of the induction and then calculate B_z with (11). The induction B_φ satisfies the boundary condition $B_\varphi(R_0) = 0$, where R_0 is the inner radius of the solenoid. In terms of the dimensionless variables $x = r/R_0$ and $y = rB_\varphi R_0^{-3/2}(\mu_0\sigma_0/(2\Delta))^{-1/2}$, Eq. (13) takes the form

$$yy' = x^2. \quad (14)$$

Its solution is

$$y = \sqrt{\frac{2}{3}(x^3 - 1)}, \quad (15)$$

or, for $h \ll \Delta$,

$$B_\varphi = \frac{1}{r}\sqrt{\frac{\mu_0\sigma_0}{3\Delta}(r^3 - R_0^3)}, \quad \delta_z = \sqrt{\frac{3\sigma_0}{\mu_0\Delta}}\frac{r}{\sqrt{r^3 - R_0^3}}. \quad (16)$$

In this limit case, the axial and azimuth components of the current density are defined as

$$\begin{aligned}B_z &= \sqrt{B_0^2 - \frac{\mu_0\sigma_0}{\Delta}(r - R_0)}, \\ \delta_\varphi &= \frac{\sigma_0}{\Delta\sqrt{B_0^2 - \mu_0\sigma_0(r - R_0)/\Delta}},\end{aligned} \quad (17)$$

where B_0 is the induction at the axis.

The outer radius R_1 of the winding is found from the condition $B_z(R_1) = 0$:

$$R_1 = R_0 + B_0^2\Delta/(\mu_0\sigma_0). \quad (18)$$

Now we can determine the stress in the insulating spacers:

$$\sigma_0 = B_0^2\frac{\Delta}{\mu_0(R_1 - R_0)} \approx \frac{B_0^2}{2\mu_0 N}, \quad (19)$$

where $N \approx (R_1 - R_0)/(2\Delta)$ is the number of turns of the winding that, it will be remembered, consists of the equithick layers that conduct the axial and azimuth currents and are separated by the insulating spacers.

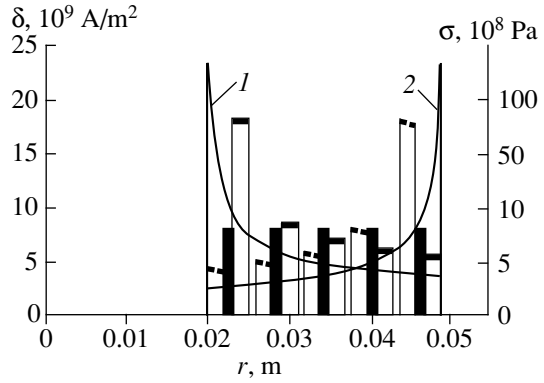


Fig. 3. Distribution of the mechanical stress and current densities in the quasi-force-free winding having five pairs of the layers: ▬ δ_z, ▪ ▪ δ_φ, and ▮ $|\sigma_r| = |\sigma_\varphi|$. (*1*, *2*) Calculation of δ_z and δ_φ by formulas (16) and (17), respectively.

As follows from approximate formula (19), the stress in our system is N times as low as the magnetic field pressure at the solenoid axis.

Figure 3 shows the numerical calculation of set (6)–(8) for of a system of five layers ($B_0 = 100$ *T*). Comparing these data with the model of continuous current distribution, we note that the approach adopted provides a reasonable approximation of the actual current distribution in the winding although the number of the layers is small. The discrepancy takes place near the boundary, where the current density can be infinitely high according to the analytic prediction. The values of the stresses ($\sigma_r = \sigma_\varphi = 7.96 \times 10^8$ Pa) are almost coincident with the calculation by formula (19).

The model of continuous current distribution can also be employed for other initial assumptions concerning the design of the winding. For example, let δ_φ be constant. It is assumed in this case that the azimuth currents are the same in each of the pairs. In Eq. (11), $B_z(r) = B_0 - \mu_0\delta_0(r - R_0)$, where R_0 is the inner radius of the solenoid and $\delta_0 = \delta_\varphi\Delta/[2(\Delta + h)] \approx \delta_\varphi/2$. At the outer boundary ($r = R_1$), $B_z(R_1) = 0$. We introduce the parameter characterizing the thickness of the winding: $g = R_0/(R_1 - R_0) = \mu_0 R_0\delta_0/B_0$. Substituting $B_z(r)$ into (11) yields the equation for the azimuth induction

$$\frac{1}{\mu_0 r}B_\varphi\frac{d}{dr}(rB_\varphi) = B_0\delta_0\left(1 - \frac{r - R_0}{B_0}\mu_0\delta_0\right). \qquad (20)$$

Then, passing to the dimensionless variables, we come to

$$xz(xz)' = x^2(1 + g) - x^3 g, \qquad (21)$$

where $x = r/R_0$ and $z = B/(B_0\sqrt{g})$.

The boundary condition for B_φ is $z = 0$ at $x = 1$. The solution has the form

$$z = \frac{1}{x}\sqrt{\frac{2(1 + g)x^3}{3} - \frac{gx^4}{2} - \frac{2}{3} - \frac{g}{6}}. \qquad (22)$$

At the outer boundary of the solenoid, we have $x = R_1/R_0 = 1 + 1/g$. The induction of the azimuth field at the outer boundary, the axial current density, and the axial current are

$$B_{\varphi e} = \frac{B_0}{\sqrt{6}(g + 1)}\sqrt{(g + 1)^4 - 4g^3 - g^4},$$

$$\delta_z = \frac{2B_0\sqrt{g}}{\mu_0 R_0}\frac{x(1 + g - gx)}{\sqrt{2(1 + g)x^3/3 - gx^4/2 - 2/3 - g/6}}, \qquad (23)$$

$$i_z = \frac{2\pi R_0 B_0}{\sqrt{6}\mu_0 g}\sqrt{(g + 1)^4 - 4g^3 - g^4}.$$

Unlike the previous example, the stresses in the spacer vary linearly in this approximation:

$$|\sigma_r| = |\sigma_z| = 2\delta_0\Delta B_0\left(1 - \frac{r - R_0}{R_1 - R_0}\right). \qquad (24)$$

They reach the maximal values at the inner boundary of the winding:

$$\sigma_{\max} = 2\Delta\delta_0 B_0 = B_0^2\Delta/[\mu_0(R_1 - R_0)] \cong B_0^2/(\mu_0 N). \qquad (25)$$

For the axial current uniformly distributed along the radius (δ_z = const), we have $B_\varphi = \mu_0\delta_0[r/2 - R^2/(2r)]$, where $\delta_z\Delta/[2(\Delta + h)] \approx \delta_z/2$. Equation (11) is reduced to the equation for the axial induction:

$$\frac{B_z}{\mu_0}\frac{dB_z}{dr} = -\frac{\mu_0\delta_0^2}{2}\left(r - \frac{R_0^2}{r}\right). \qquad (26)$$

In terms of the dimensionless variables $x = r/R_0$ and $y = \sqrt{2}\,B_z/(\mu_0\delta_0 R_0)$, (26) becomes

$$yy' = -x + 1/x. \qquad (27)$$

The boundary condition is $y = 1$ at $x = 1$. At the outer boundary, $y = 0$. The solution has the form

$$y = \sqrt{2 + 2\ln x - x^2}. \qquad (28)$$

The outer radius of the winding is found from the condition $y = 0$, to which the value $x_1 = R_1/R_0 \approx 1.77$ corresponds. The stresses in the spacer, $|\sigma_z| = |\sigma_\varphi| = \delta_z\Delta B_\varphi \approx \delta_0^2\Delta\mu_0 R_0(x - 1/x)$, reach the maximum at the outer boundary:

$$|\sigma|_{\max} \cong 2.4(\Delta/R_0)B_0^2/\mu_0 \cong 0.95B_0^2/(\mu_0 N). \qquad (29)$$

Here, we used the approximate equality

$$N \approx (R_1 - R_0)/(2\Delta) = 0.385R_0/\Delta,$$

which is valid if $d \gg \Delta$ and h. In this approximation, the azimuth current density is given by

$$\delta_\varphi = -\frac{2B_0}{\mu_0 R_0}\frac{1 - x^2}{x\sqrt{2 + 2\ln x - x^2}}. \qquad (30)$$

The value of δ_φ infinitely grows at the edge of the winding ($r = R_1$), but the product $\delta_\varphi B_z$ remains finite.

MAGNETIC SYSTEM HAVING A LARGE NUMBER OF LAYERS WITH DIFFERENT CURRENT DIRECTIONS

If the current distribution in the cylindrical region is force-free, the current density vector continuously varies along the radius. Such a current distribution can be approximated by a set of current layers where the current direction discretely varies [4]. It has been shown [8] that the stresses in a thin-walled cylinder sharply decrease if the current makes an angle of about $\pi/4$ with the cylinder axis. Also, the volume electromagnetic forces in a multilayer system approximating the force-free system can be drastically reduced according to [7]. However, in [7], the balance problem for each of the layers has not been discussed although the equilibration of the layers is necessary for effectively reducing elastic stresses.

Let us introduce the coordinate x that is reckoned from the inner boundary of the nth layer, $x = r - r_n$ (Fig. 4). In this layer, the axial component of the current density is x-independent and equals $\delta_{z,n}$, while the azimuth component is given by (3); that is, $\delta_\varphi = \delta_{\varphi,n} r_n/(r_n + x)$. Up to the terms of the order of x/r_n, the following expression for the volume force is valid:

$$f_r(x) \approx \delta_{\varphi,n} B_z(r_n)(1 - x/r_n) - \delta_{z,n} B_\varphi(r_n)(1 - x/r_n) - \mu_0 x(\delta_{\varphi,n}^2 + \delta_{z,n}^2). \quad (31)$$

The resultant of the volume forces at a point x is

$$F_{r,n}(x) = \delta_{\varphi,n} B_z(r_n) x\left(1 - \frac{x}{2r_n}\right) - \delta_{z,n} B_\varphi(r_n) x\left(1 - \frac{x}{2r_n}\right) - \frac{\mu_0 x^2}{2}(\delta_{\varphi,n}^2 + \delta_{z,n}^2) = a_n x + b_n \frac{x^2}{2}, \quad (32)$$

where $a_n = \delta_{\varphi,n} B_z(r_n) - \delta_{z,n} B_\varphi(r_n)$ and $b_n = -a_n/r_n - \mu_0(\delta_{\varphi,n}^2 + \delta_{z,n}^2)$.

For this case, Eq. (6) for the induction components remains valid. It can be obtained from the balance condition for the layers $F_{r,n}(\Delta) = 0$.

The essential difference of the system under study, where the axial and azimuth currents are combined in one layer, from that with paired current layers is that the components $\delta_\varphi B_z$ and $(-\delta_z B_\varphi)$ partially cancel. The volume force in a layer is alternating-sign in this case, and its integral $F_{r,n}(x)$ reaches a maximum at the point $x = \Delta/2$. The value of $F_{r,n}(\Delta/2)$ is given by two equivalent expressions:

$$F_{\max n} = (\delta_{\varphi,n} B_z(r_n) - \delta_{z,n} B_\varphi(r_n))\frac{\Delta}{4}, \quad (33.1)$$

$$F_{\max n} = \frac{\mu_0 \Delta^2}{8}\left(\delta_{z,n}^2 + \delta_{\varphi,n}^2 + \frac{\delta_{\varphi,n} B_z(r_n) - \delta_{z,n} B_\varphi(r_n)}{\mu_0 r_n}\right). \quad (33.2)$$

It can be shown that $|F_{\max n}|$ is the estimate of the modulus of the stresses in the layers. Note that in the system being considered, the stresses are markedly lower than in the system with paired layers. The former approximates a force-free system more closely, since the current components are not spaced. Because of the partial cancellation of the volume forces resulting from the poloidal and toroidal fields, the stresses decrease with the number of layers more drastically than in the paired-layers system. This follows from the comparison of Eqs. (33) and (9). The current densities in the layers are of the order of $B_0/(\mu_0 \Delta N)$, where B_0 is the induction at the axis. Hence, according to (33.2), $F_{\max n}$ has the order of $B_0^2/(\mu_0 N^2)$.

For a given distribution of the volume electromagnetic forces in the layer, the stresses can be calculated with the well-known formulas from the theory of elasticity (see, e.g., [9]):

$$\sigma_{r,\varphi} = -\frac{1+\Theta}{2}\int_{r_n}^{r} f(r)dr \mp \frac{1-\Theta}{2r^2}\int_{r_n}^{r} f(r) r^2 dr + C_1 \mp \frac{C_2}{r^2}, \quad (34)$$

where $f(r)$ is the radial volume force, $\Theta = \mu/(1-\mu)$, μ is Poisson's ratio, and C_1 and C_2 are constants defined by

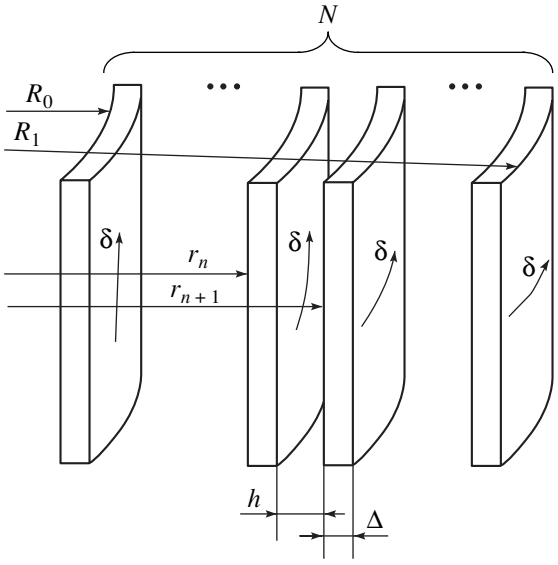


Fig. 4. Quasi-force-free winding with variously directed currents.

the boundary conditions $\sigma_r(r_n) = 0$ and $\sigma_r(r_n + \Delta) = 0$.

Using formula (31) for $f(r)$, we expand expression (34) into the series in the small parameter x/r_n, where $x = r - r_n$, rejecting the terms of the order of x^3/r_n^2 or higher:

$$\sigma_{r,\varphi} \approx -\frac{1+\Theta}{2}\left(a_n x + \frac{b_n x^2}{2}\right) \mp \frac{1-\Theta}{2} \times \left[a_n x - \left(\frac{a_n}{r_n} - \frac{b_n}{2}\right)x^2\right] + C_1 \mp \frac{C_2}{r_n^2}\left(1 - \frac{2x}{r_n}\right), \quad (35)$$

where a_n and b_n are defined above.

The constants C_1 and C_2 are found from conditions $\sigma_r(0) = 0$ and $\sigma_r(\Delta) = 0$:

$$C_1 \cong -(1-\Theta)a_n\frac{\Delta}{2}\left(1 + \frac{3}{2}\frac{\Delta}{r_n}\right), \quad C_2 = C_1 r_n^2. \quad (36)$$

In view of the balance condition for the layer $a_n\Delta + b_n\Delta^2/2 = 0$, we come to the expressions for the voltages (the terms of the order of Δ^3/r_n^2 are rejected):

$$\sigma_r \approx -a_n\left(x - \frac{x^2}{\Delta}\right) + (1-\Theta)\frac{a_n x}{2r_n}(x - \Delta),$$

$$\sigma_\varphi = -a_n\Theta\left(x - \frac{x^2}{\Delta}\right) \quad (37)$$

$$-(1-\Theta)a_n\left[\frac{x^2}{2r_n} + \frac{\Delta}{2}\left(2 - \frac{2x}{r_n} + \frac{3\Delta}{r_n}\right)\right].$$

The moduli of both voltages are maximal at the midpoint of the layer ($x = \Delta/2$):

$$|\sigma_r|_{\max} = F_{\max n}\left(1 + (1-\Theta)\frac{\Delta}{8r_n}\right),$$
$$|\sigma_\varphi|_{\max} = 4F_{\max n}\left(1 - \frac{3\Theta}{4} + (1-\Theta)\frac{9}{8}\frac{\Delta}{r_n}\right). \quad (38)$$

Both components of the stress tensor coincide with the maximal resultant of the volume forces $(F_n)_{\max} = a_n\Delta/4$, up to the close-to-unity factors.

As an example, let us consider a small-curvature winding, for which the terms of the order of Δ/r_n can be neglected. In essence, this is a set of N plane conductive layers of thickness Δ. For simplicity, we assume that the thickness of the insulating spacers between the layers is negligibly small. In the limit $N \longrightarrow \infty$, we have the force-free field for which $|B| = \sqrt{B_z^2 + B_\varphi^2} = \text{const}$. With the balance condition for the set of the discrete layers met, the currents $\delta_{z,n}$ and $\delta_{\varphi,n}$ are expressed through the inductions on both sides as $\delta_{\varphi,n} = -(B_{z,n+1} - B_{z,n})/(\mu_0\Delta)$ and $\delta_{z,n} = (B_{\varphi,n+1} - B_{\varphi,n})/(\mu_0\Delta)$; then, the expressions for the maximal resultants take the form

$$(F_n)_{\max} = \frac{1}{4\mu_0}[B_{z,n}^2 + B_{\varphi,n}^2 - (B_{z,n}B_{z,n+1} + B_{\varphi,n}B_{\varphi,n+1})]. \quad (39)$$

At the outer boundary, $B_{z,1} = B_0$ and $B_{\varphi,1} = 0$; at the inner one, $B_{z,N+1} = 0$ and $B_{\varphi,N+1} = B_0$, where N is the number of the layers.

A discrete system where the induction between the layers has the same modulus but different directions is an analog of a force-free one. An example is the distribution

$$B_{z,n} = B_0\cos\frac{(n-1)\Delta\pi}{2d},$$
$$B_{\varphi,n} = B_0\sin\frac{(n-1)\Delta\pi}{2d}, \quad (40)$$

where $d \approx N\Delta$ is the thickness of the winding. Here,

$$(F_n)_{\max} = \frac{B_0^2}{4\mu_0}\left(1 - \cos\frac{\Delta\pi}{2d}\right), \quad (41)$$

that is, the resultants are the same in all of the layers and the winding is equiloaded. For each of the layers, the maximal volume force

$$f_{\max} = \frac{B_0^2}{\mu_0\Delta}\left(1 - \cos\frac{\Delta\pi}{2d}\right) \quad (42)$$

is observed at its inner boundary ($x = 0$). For the single-layer case, $\Delta = d$ and $F_{\max} = 0.5B_0^2/(2\mu_0)$. This case has been analyzed in [8]. For two and three layers, we have, respectively, $F_{\max} \approx 0.15B_0^2/(2\mu_0)$ and $F_{\max} \approx 0.065B_0^2/(2\mu_0)$. From these estimates, we can infer that even simple two- and three-layer systems with variously directed currents may stress a layer much weaker than the pressure due to a magnetic field of induction B_0. In the single-layer example, the current density vector makes an angle of $\pi/4$ with the z axis; in the double- and treble-layer systems, the direction of this vector changes by the same angle from layer to layer. The same picture is observed for the general case of N layers. For $N \gg 1$, we have $\Delta\pi/(2d) = \pi/(2N) \ll 1$; hence,

$$F_{\max} \approx \frac{B_0^2}{2\mu_0}\frac{\pi^2}{16N^2}, \quad f_{\max} \approx \frac{B_0^2\pi^2}{8\mu_0\Delta N^2}. \quad (43)$$

The order of $F_{\max}$ agrees with the above estimate $F_{\max} = \text{const}/N^2$.

This result is corroborated by numerical calculations for real systems (without the small-curvature assumption). Figure 5 demonstrates the results of computer simulation for an equiloaded five-layer winding with cylindrical layers. A comparison with Fig. 3 shows

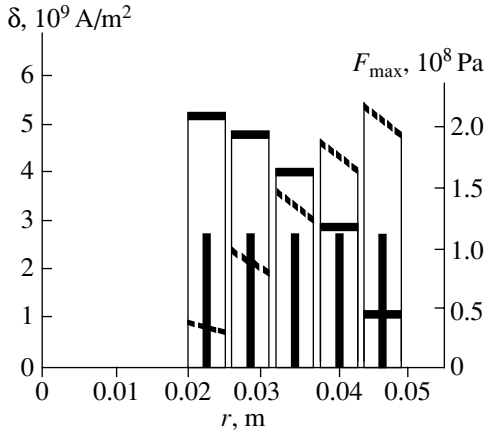


Fig. 5. Distribution of the resultant volume force and the current densities in the quasi-force-free winding having five layers with inclined currents (numerical calculation): ▬ σ_z, ▬ ▬ σ_φ, and ▮ E_{max}.

that the stresses in the layers with variously directed currents are roughly seven times as low as those in the system of paired current layers, all other parameters being nearly the same. For $B_0 = 100$ T, the maximum of the resultants F_{max} is 1.1×10^8 Pa. Note that the calculation by formula (41), which ignores the curvature of the layers, gives a similar result ($F_{max} = 0.97 \times 10^8$ Pa).

CONCLUSIONS

(1) In systems with thin cylindrical current layers, the electromagnetic forces can be drastically reduced if the layers are split into pairs of conductors with orthogonal (azimuth and axial) currents and the forces acting on either conductor are counterbalanced. The residual stresses in the separating spacer are of the order of $B_0^2/(\mu_0 N)$, where B_0 is the induction at the axis and N is the number of the layers.

(2) The decrease in the residual stresses is still greater in systems of balanced cylindrical layers with variously directed currents. Here, the stresses may be of the order of $B_0^2/(\mu_0 N^2)$ provided that the current distribution over the layers is appropriated selected.

REFERENCES

1. G. A. Shneerson, Zh. Tekh. Fiz. **56** (1), 36 (1986) [Sov. Phys. Tech. Phys. **31**, 20 (1986)].
2. G. A. Shneerson, IEEE Trans. Magn. **28**, 505 (1992).
3. G. A. Shneerson, in *Megagauss Magnetic Field Generation and Pulsed Power Applications*, Ed. by M. Cowan and R. B. Spielman (Nova Science, New York, 1994), Vol. 1, pp. 25–34; 171–172.
4. G. A. Shneerson, V. Yu. Khosikov, and E. L. Amromin, Plasma Devices Op. **4**, 321 (1998).
5. V. A. Ignatchenko and M. M. Karpeko, Zh. Tekh. Fiz. **38**, 200 (1968) [Sov. Phys. Tech. Phys. **13**, 143 (1968)].
6. H. J. Shneider-Muntau and P. Rub, The Hilex Project. Colloques Internatiaux C.N.R.S. No. 242. Physique Sous Champes Intenses (1974), p. 161.
7. S. G. Bodrov and G. A. Shneerson, Pis'ma Zh. Tekh. Fiz. **20** (9), 47 (1994) [Tech. Phys. Lett. **20**, 361 (1994)].
8. A. A. Kuznetsov, Zh. Tekh. Fiz. **31**, 650 (1961) [Sov. Phys. Tech. Phys. **6**, 472 (1961)].
9. D. B. Montgomery, *Solenoid Magnet Design* (Wiley, New York, 1969; Mir, Moscow, 1971).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 114–120. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 117–124.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

The Diagnostics of Thermal Kinetic Coefficients for the Optimization of Film Bolometer Properties

Yu. V. Medvedev*, Yu. M. Nikolaenko*, A. M. Grishin, and S. I. Khartsev****

** Galkin Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 83114 Ukraine*

e-mail: nik@kinetic.ac.donetsk.ua

*** Royal Institute of Technology, S-100 44, Stockholm, Sweden*

Received September 13, 2000; in final form, May 16, 2001

Abstract—A pulsed microwave technique for the fast and accurate diagnostics of a complete set of thermal kinetic coefficients in single-layer film structures is developed. The technique is tested on $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3$ films on $LaAlO_3$ and $SrTiO_3$ substrates—the structures featuring a temperature resistance coefficient of 5–7%, which is the highest for manganite-containing materials at room temperature. The results of numerical simulation show that the performance of the uncooled bolometers can be improved by optimizing the thermal conductivity of the substrate and matching the thermal resistances of the film–substrate and substrate–thermostat interfaces. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The recent advances in the film technology of high-temperature superconductors and manganite-containing materials have been related largely to the optimization of the electrophysical properties of the films. Other parameters, such as those responsible for heat exchange in the system, have been overlooked. Therefore, it is not surprising that the film structures that are the best in electrical properties do not have thermodynamic parameters optimal in terms of bolometric applications [1, 2]. This shows up in the pronounced nonlinearity of the amplitude–frequency characteristic of the bolometers and also in the large maximal time constant of the response. The potentialities for simultaneously improving the operating speed and the power sensitivity of the bolometers by optimizing the shape and size of the film or the entire structure have been practically exhausted. For cooled bolometers, this problem has been partially solved with hybrid designs [3, 4]. However, even for the uncooled devices, their performance can be improved directly by optimizing the thermal resistance of the film–substrate interface: $R_{fs} = (S\lambda_{fs})^{-1}$, where S is the surface area of the film and λ_{fs} is the thermal conductivity of the film–substrate interface. Although this conceptually simple idea has long been known [3], its implementation requires advanced film-deposition processes that provide desired values of R_{fs} to be developed. Moreover, simple and yet reliable techniques for measuring this parameter should be elaborated.

In one stationary method [5], R_{fs} is determined by measuring the temperature difference between two metal strips that act both as heaters and temperature-sensitive elements. In the case of low currents, either transducer measures the temperature of the film that equals that of the substrate. As the current through one of the strips grows, the transducer readings give the temperature drop, which depends on R_{fs}. The disadvantage of the method is the need for the additional interface between the film and the transducer and also that the temperature gradient in the substrate cannot reliably be measured.

Another, nonstationary method for R_{fs} evaluation measures the time constant of an optical bolometric response [5, 6]. This method is more powerful and gains more reliable information on the thermal coefficients of the structure, but its potentialities for evaluating small λ_{fs}, $\lambda_{fs} \le 10^5$ W/(m^2 K), are very limited when the simplified approach to the interpretation of the heat distribution in the system is applied. As for high values of λ_{fs}, $\lambda_{fs} \sim 10^7$ W/(m^2 K), they are difficult to measure, because the nonbolometric part of the response may far exceed the bolometric part.

The nonstationary method of R_{fs} diagnostics that operates with both the temporal and amplitude characteristics of the bolometer response [7] is free of the above disadvantages. In this method, the film is heated by microwave pulses and the time-dependent temperature difference $\Delta T_f(t)$ between the film and thermostat temperatures (T_f and T_0, respectively) is recorded and analyzed. The thermal kinetic coefficients (TKCs) are derived by numerically simulating the function $\Delta T_f(t)$ with regard for information borne by its different sections. The TKC orders can also be estimated with approximate analytical expressions for temperature drops ΔT_i and time constants τ_i (i = 1–3) that characterize stationary heat transfer (1) at the film–substrate interface, (2) through the substrate, and (3) at the substrate–thermostat interface. Each of the time instants τ_i corresponds to the related section of the $\Delta T_f(t)$ curve

and is found by irradiating the films with pulses of appropriate duration.

In this work, the method developed in [7] is generalized for $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/LaAlO_3$ and $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/SrTiO_3$ structures with "dead" layers of different thicknesses and at near-room temperatures. We succeeded in determining the complete set of the TKCs, including R_{fs}. The procedure allows the finding of desired parameters even if the relaxation times of the thermal flow through the film–substrate interface and through the substrate become comparable. Of great importance is that the exact and approximate expressions that describe the distribution of the energy absorbed in the system give similar values of desired parameters.

From the results of simulation, we found the correlation between the basic parameters and the TKCs of the bolometer. It turned out that the TKC optimization can raise the operating speed of the device by three orders of magnitude without degrading its power sensitivity. This provides fresh insight into the relation between the bolometric response, the physical properties of the film systems, and the structures of the interfaces.

MEASUREMENTS MADE BY THE TEMPORAL THERMOMETRY METHOD

The temporal thermometry method, which characterizes heat transfer by phonons through interfaces, was extended for $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/LaAlO_3$ and $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/SrTiO_3$ magnetoresistive film structures (hereafter, structures 1 and 2). The films 0.5 μm thick were applied by pulsed laser deposition [1]. The substrates measured $3 \times 6.7 \times 0.6$ and $2 \times 5 \times 0.5$ mm for structures 1 and 2, respectively. The composition of the films provided the maximal value of the temperature resistance coefficient (TRC) at a level $\partial(\ln(R))/\partial T$ = 5–7% at room temperature. The temperature dependence of the resistance of the films that is normalized to its maximum (1875 and 9990 Ω for structures 1 and 2, respectively) is shown in Fig. 1.

The measuring device consisted of a copper waveguide flange (thermostat) placed in the vacuum chamber. The film structure was glued to a mica plate on the substrate. The plate covered the 3.6×7.2-mm waveguide window. The temperature of the flange was maintained with the copper thermostat and controlled by an electronic unit with an accuracy of 10^{-2} K. The heat was removed from the thermostat to the vacuum chamber body via a copper wire with a thermal resistance of about 10 K/W.

The time dependence of the film temperature was recorded with a 10-bit ADC with a maximal speed of 10^7 samples per second. The temperature was determined from the value and the sign of change of the resistance. Upon recording the time dependence of the temperature, the microwave power P_{mw} was applied in two ways. For the larger time scale, a single pulse was applied. For shorter times, a train of the pulses P_{mw} with a pulse period-to-pulse duration ratio of 2 was used. In the latter case, the film is periodically heated and cooled. Such an approach makes it possible to study any portion of the $T_f(t)$ curve by controlling the pulse duration, detector sensitivity, and thermostat temperature.

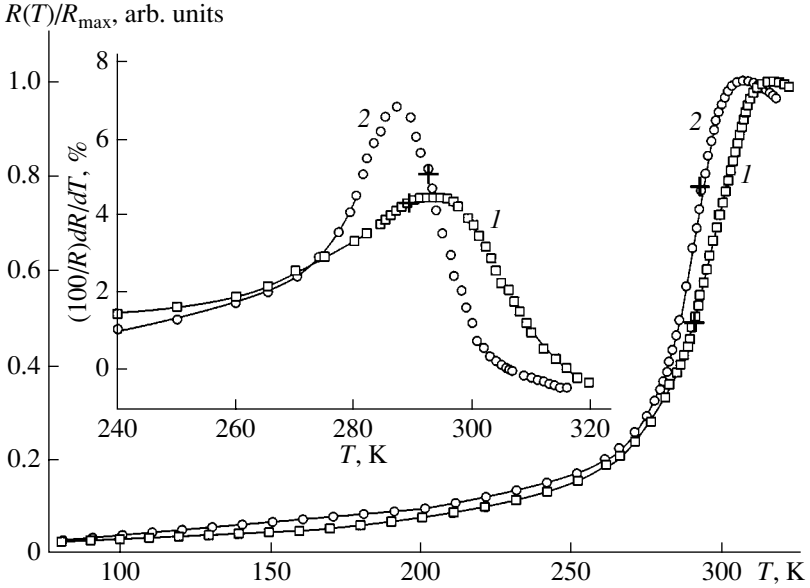


Fig. 1. Temperature dependence of the film resistance normalized to its maximum. The insert shows the temperature dependence of the temperature resistance coefficient. (*1*) $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/LaAlO_3$ and (*2*) $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/SrTiO_3$ structures.

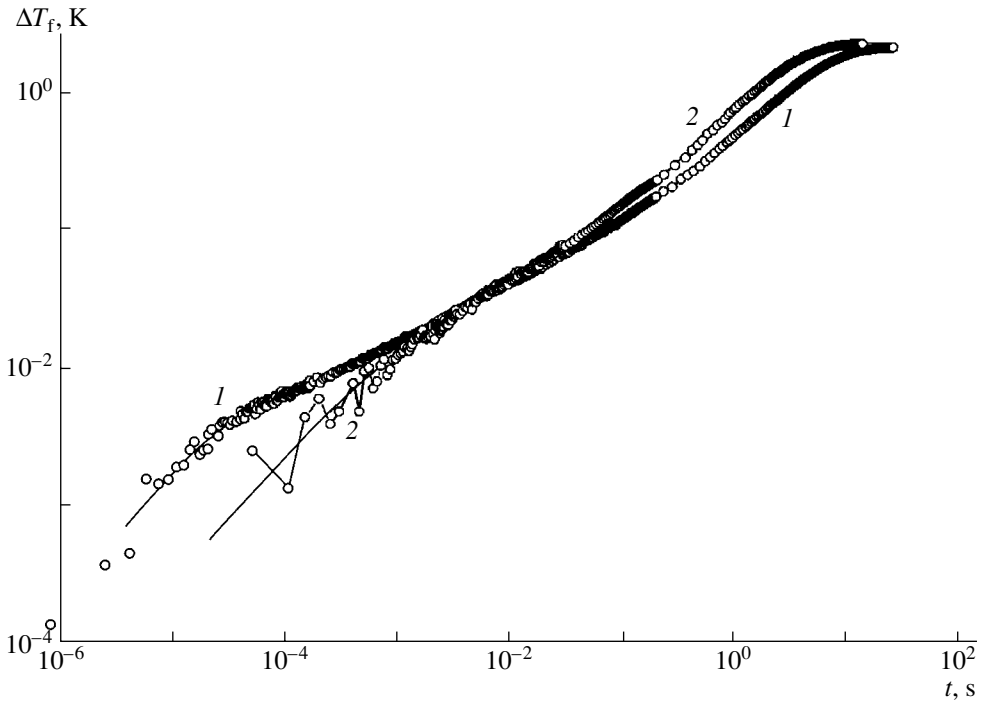


Fig. 2. Time dependences of the heating temperature for (*1*) $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/LaAlO_3$ and (*2*) $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/SrTiO_3$ structures in the log–log scale. Circles, data points; curves, calculation.

Figure 2 demonstrates experimental $\Delta T_f(t)$ values (circles) that were obtained by summing several files containing thermometric data for various time scales. The solid curves are the result of calculation using the parameters providing the best fit to the experimental curves. In Fig. 3, the curves are drawn in the linear time scale to illustrate the degree of coincidence between the theory and the experiment.

Mathematically, the analysis of the temperature dependence of the film temperature and the calculation of the thermodynamic coefficients of the structures rely on the set of equations [7–10]

$$c_f\rho_f V_f(\partial T_f/\partial t) = -(S\lambda_{fs})(T_f - T_s(x = 0, t)) + P_J, \quad (1)$$

$$c_s\rho_s(\partial T_s/\partial t) = \lambda_s(\partial^2 T_s/\partial x^2). \quad (2)$$

Here, c_f and c_s are the specific heat capacities of the film and the substrate, respectively; ρ_f and ρ_s are the densities of the film and the substrate, respectively; S is the surface area of the film; λ_{fs} is the thermal conductivity of the film–substrate interface; P_J is the Joule heat; λ_s is the thermal conductivity of the substrate; and V_f is the volume of the film. The initial and boundary conditions are written as

$$T_f = T_s = T_0, \quad t = 0, \quad (3)$$

$$\lambda_s(\partial T_s/\partial x) = \lambda_{fs}(T_f - T_s), \quad x = 0, \quad (4)$$

$$\lambda_s(\partial T_s/\partial x) = \lambda_{st}(T_s - T_0), \quad x = D. \quad (5)$$

The point $x = 0$ is assumed to lie on the film–substrate interface (the x axis is perpendicular to the film surface). The point $x = D$ lies on the substrate–thermostat interface. The temperature gradient in the film is ignored. The thermal conductivities of the interfaces, $\lambda_{fs} = (SR_{fs})^{-1}$ and $\lambda_{st} = (SR_{st})^{-1}$ (R_{st} is the thermal resistance of the substrate–thermostat interface), are assumed to be T-independent; that is, we believe that the temperature change is small: $\Delta T_f(t) \ll T_0$. Also, we do not consider instability effects due to film self-heating [3, 9] when $P_J(T)$ sharply varies with T.

It follows from Figs. 2 and 3 that the process reaches the nonequilibrium steady state in several stages having different rates of change of $\Delta T_f(t)$. To better understand the qualitative relation between the run of the $\Delta T_f(t)$ curve and the process of heat transfer, we invoke idealized models, which are usually used in the analysis of the bolometric response (see also [7]). Once the time-invariable power P_J (Joule heat) has been applied, the time behavior of the film temperature within the model "substrate–perfect thermostat ($c_s = \infty$) is given by the exponential expression

$$\Delta T_f(t) = \Delta T_1(1 - \exp(-t/\tau_1)), \quad (6)$$

where $\tau_1 = c_f\rho_f V_f R_{fs}$ is the time constant and $\Delta T_1 = P_J R_{fs}$ is the temperature amplitude at the film–substrate interface.

When a constant thermal flow propagates in the substrate, the temperature at the substrate boundary facing

the film is a composite function including the linear and saturating parts, which describe, respectively, an increase in the mean substrate temperature and the propagation of the flow between the interfaces [10]:

$$T_s(x = 0, t) = T_0 + \frac{P_J}{c_s\rho_s V_s} + \Delta T_2\left(1 - \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\exp\left(-n^2\pi^2\frac{t}{\tau_2}\right)\right). \quad (7)$$

Here, V_s is the substrate volume, $\tau_2 = c_s\rho_s V_s/(S\lambda_s)$, and $\Delta T_2 = P_J D/(3S\lambda_s)$ is the temperature drop between the upper and lower boundaries of the substrate. Finally, if the substrate–thermostat interface is viewed as a whole structure, the dependence is again exponential:

$$\Delta T_f(t) = \Delta T_3(1 - \exp(-t/\tau_3)). \quad (8)$$

In this case, the time constant $\tau_3 = (c_f\rho_f V_f + c_s\rho_s V_s)R_{st}$ corresponds to the stationary heat transfer due to the temperature drop $\Delta T_3 = P_J R_{st}$ between the film and the thermostat. Approximate relationships (6)–(8) clearly demonstrate the qualitative relation between the TKCs of the structure and the behavior of the $\Delta T_f(t)$ curve. If $\tau_1 \gg \tau_2 \gg \tau_3$, they can be used to estimate the coefficients on the order of magnitude.

The Joule heat is the key parameter that enables the quantitative evaluation of the temperature amplitudes and the TKCs. There are a number of reasons why P_J cannot be estimated from the power P_{mw} applied. We developed an original procedure of P_J normalization that does not depend on film irradiation conditions. The value of P_J was normalized using the dependence $\Delta T_J(P_J)$, where $P_J = I^2R$, R is the film resistance, and I is the bias current. This dependence was also used to make sure that the film temperature determined from the resistive response is correct. The associated plot must be and is linear (Fig. 4). This dependence is obtained by converting the dependence $R(I)$ in view of the function $R(T)$ derived at a small bias current. The weak nonlinearity of $R(P_J)$ (the lower insert in Fig. 4) results from the nonlinearity of $R(T)$ and can serve as a criterion for determining the smallness of ΔT.

In Fig. 4, closed circles and squares show the $\Delta T_f(P_J)$ dependence obtained in the small-current regime when the film was heated by microwave pulses of variable power. The value of P_{mw} was set with an attenuator. Due to the low resistance of the electrical contacts (Ag strips deposited on the edges of the films), the direct-detection effect was completely absent even with the two-point probe measurements. From the linearity of two $\Delta T_f(P_J)$ dependences, it follows that the Joule heat is proportional to the microwave power and that the normalization of P_J may be accomplished by multiplying P_{mw} by a factor that provides the closest fit between the two linear functions (Fig. 4). The upper insert in Fig. 4 gives the values of the total thermal

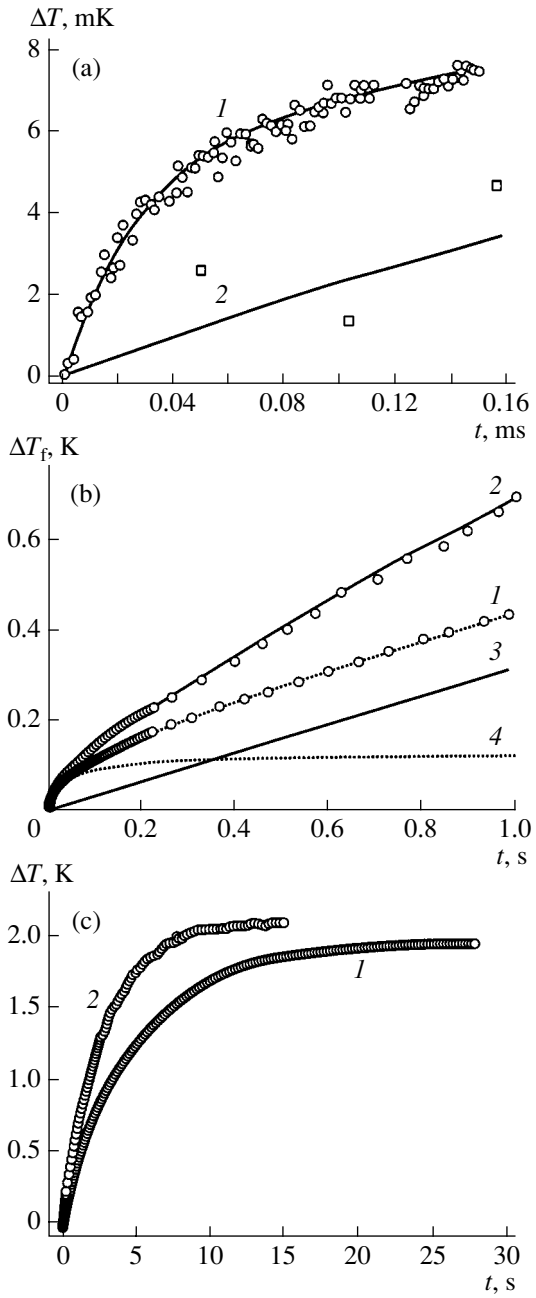


Fig. 3. The same as in Fig. 2 in the linear time scale. (*3*) Linear and (*4*) saturating parts of curve *1*.

resistance of the film structure, $R_{tot} = \Delta T/P_J$, calculated from the above $\Delta T_f(P_J)$ dependences.

The normalized power and the values of the TKCs obtained after fitting the curves for structures 1 and 2 (Fig. 2) are listed in the second and fourth rows of the table. In the first and third rows, typical literature data [11–14] for the coefficients that were used in fitting are presented. Also, the first and third rows list the time constants τ_i and temperature amplitudes ΔT_i calculated with (6)–(8).

Comparing the literature data for the TKCs (the first and third rows) with the results of fitting (the second

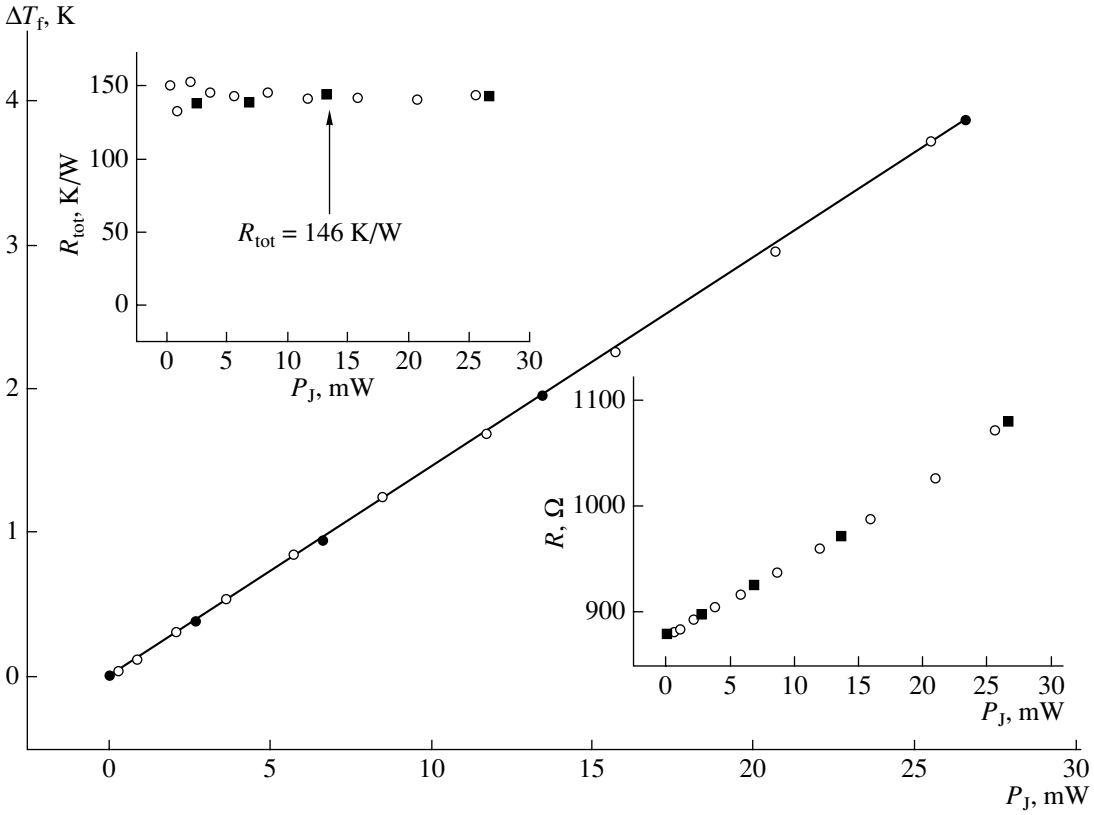


Fig. 4. Heating temperature vs. Joule heat for film 1. Lower insert: resistance of film 1 vs. Joule heat; upper insert: total thermal resistance of the entire film–thermostat system vs. Joule heat for structure 1. ○, dc heating; ● and ■, microwave heating. The microwave power level (P/P_{max} = 1, 0.5, 0.25, 0.1, and 0) was set with an attenuator.

and fourth rows) indicates that some of the coefficients greatly diverge. In particular, λ_{fs} for structure 1 turned out to be two orders of magnitude higher. This is because the thin dead layer is present on the substrate surface [1, 15]. This layer is of low electrical and, as was revealed later, thermal conductivity. The low thermal conductivity (λ_s) of the $LaAlO_3$ substrate can be associated with its mosaic structure.

The discrepancy between the literature and experimental specific heat capacities of the $SrTiO_3$ substrate (the latter value is roughly half as high as the former) is more difficult to explain. Moreover, this substrate is of low density (4760 versus 6450 kg/m^3). We, therefore, report the data for structure 2 merely to demonstrate the applicability of the method to nonstandard substrates. Note, however, that the thermal conductivity λ_s of the $SrTiO_3$ substrate agrees with data in [14] but is 4.3 times lower than in [11].

One more feature of structure 2 is the much higher noise level compared with structure 1. Because of this, we failed to take thermometric measurements from the film in the time scale of the order of τ_1. Therefore, curve *2* in Fig. 2 was fitted with the help of C_f for film 1 with regard for the size of film 2. Note that λ_{fs} for structure 2 is estimated at a level of 10^7 W/(m^2 K) or higher.

In our opinion, the difference between the values of λ_{fs} for structure 1 and 2 arises because of the different film–substrate mismatches. According to [1], the dead layer in structure 2 should be one order of magnitude thinner than in structure 1.

IMPROVEMENT OF THE TIME CHARACTERISTICS OF THE BOLOMETRIC RESPONSE FROM FILM STRUCTURES

In view of good agreement between the experimental and analytic $T_f(t)$ dependences in the film structures with the parallel heat flux in film–substrate–thermostat system, it is easy to find ways for improving the time characteristics of uncooled bolometers. It should be noted that a general optimization procedure for the film structures is still lacking [4].

An increase in the thermal resistance of the film–substrate interface seems to be the most promising way in this respect. Since the specific heat capacity of the film is three or more orders of magnitude lower than that of the substrate, one might expect that the temperature amplitude ΔT_1 will become comparable to ΔT_3 as R_{fs} rises to R_{st}. In this case, the time constant τ_1 will be several orders of magnitude smaller than τ_3. The shape

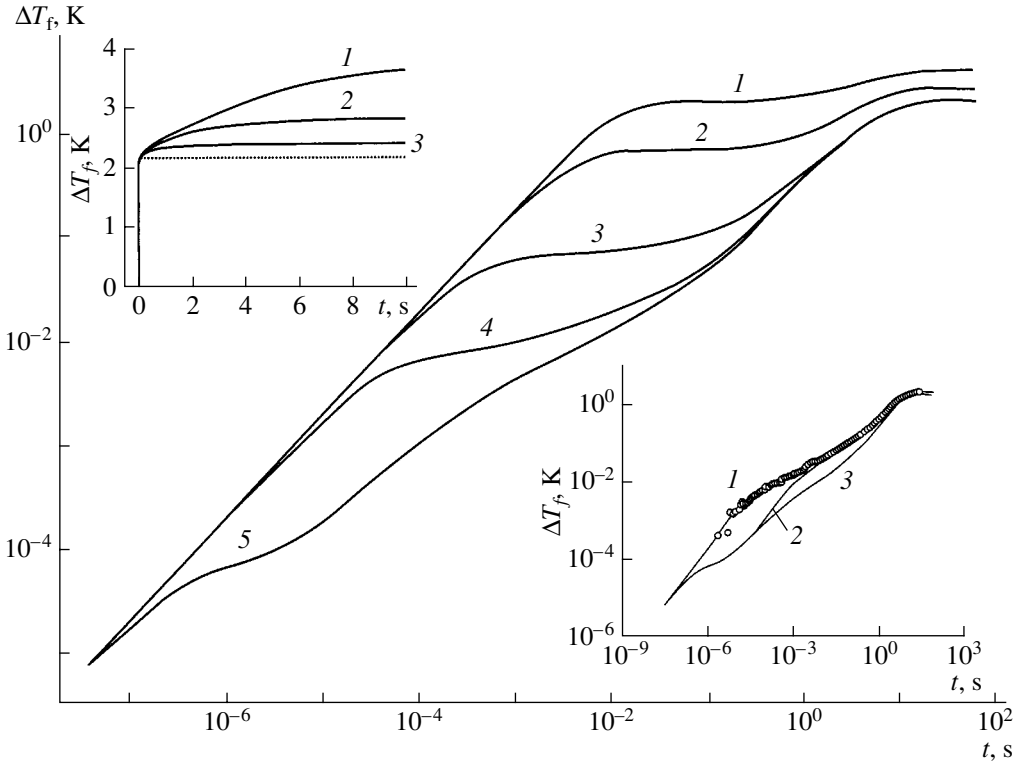


Fig. 5. Variation of the film temperature amplitude with thermal conductivity of the film–substrate interface. λ_{fs} = (*1*) 3×10^2, (*2*) 10^3, (*3*) 10^4, (*4*) 10^5, and (*5*) 10^7 W/(m^2 K). The upper insert shows the same dependence for the substrate–thermostat interface. λ_{st} = (*1*) 300, (*2*) 1000, and (*3*) 3000 W/(m^2 K) at λ_{fs} = 300 W/(m^2 K). The other parameters are given in the table (first row). The lower insert shows the calculation of the film temperature using the parameters listed in the first row of the table (curve *3*). For curve *1*, λ_s = 1.4 W/(m K) and $\lambda_{fs} = 10^5$ W/(m^2 K); for curve *2*, λ_s = 1.4 W/(m K).

of the time dependence of the film temperature as a function of R_{fs} is demonstrated in Fig. 5 in the log–log scale.

The curves are obtained for the typical parameters of the film and the substrate (the first row in the table). Curves *1–4* are plotted for the thermal resistance R_{fs}, which is, respectively, 30000, 10 000, 1000, and 100 times higher than R_{fs} for curve *5*. For curve *1*, R_{fs} = 147 K/W, i.e., roughly equals R_{st}. In this case, the bolometer has the response twice as high as that at the zero modulation frequency, and its response time is three orders of magnitude smaller.

Another, perhaps less significant, possibility of improving the bolometer performance is to decrease the thermal conductivity of the substrate (see the lower insert in Fig. 5). Curve *3* here corresponds to the standard parameters of the film and the structure. For curve *2*, the thermal conductivity of the substrate is reduced from 10.7 to 1.4 W/(m K); for curve *1*, along with this reduction, the thermal conductivity of the

Thermal kinetic coefficients, temperature amplitudes, and time constants for the $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/LaAlO_3$ and $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3SrTiO_3$ structures

No.	Substrate material	C_f, J/K	λ_{fs}, W/(m^2 K)	c_s, J/kg K	λ_s, W/m K	λ_{st}, W/(m^2 K)	τ_1, μs	τ_2, s	τ_3, s	ΔT_1, mK	ΔT_2, K	ΔT_3, K	P_J, mW
1	$LaAlO_3$	4×10^{-5}	1×10^7	441	10.7	335	0.2	0.11	5.9	0.065	0.011	1.9	13
2	$LaAlO_3$	6.43×10^{-5}	1×10^5	483	1.4	380	–	–	–	–	–	–	13
3	$SrTiO_3$	1.1×10^{-5}	1×10^7	830	2.46	500	0.11	0.415	3.95	0.098	0.066	1.95	9.75
4	$SrTiO_3$	–	1×10^7	437	2.3	503	–	–	–	–	–	–	9.75

Note: The first and third rows list the kinetic coefficients calculated from literature data [11–15] and used as the input for the fitting procedure. The temperature amplitudes and the time constants are calculated from model formulas (6)–(8). The second and fourth rows list the parameters obtained by fitting.

film–substrate interface is reduced from 10^7 to 10^5 W/(m^2 K). As a result, the temperature amplitude of the response increased three to ten times in the range 10^{-3}–10^{-1} s for curve *2* and in the range 10^{-6}–10^{-1} s for curve *1*. Note that curve *1* corresponds to the $La_{0.7}(Pb_{0.7}Sr_{0.3})_{0.3}MnO_3/LaAlO_3$ structure. Thus, the bolometer performance can be tuned to specific applications by varying the thermal conductivity of the substrate. However, the effect of varying λ_s is much weaker than when the thermal resistance of the film–substrate interface varies.

The disadvantage of the function $\Delta T_f(t)$ in Fig. 5 is the relatively large amplitude of the "slow" response ΔT_3. This, in particular, makes the amplitude–frequency characteristic nonlinear and causes signal distortions when power pulses are applied. The amplitude–frequency characteristic can be improved by matching the thermal conductivities λ_{fs} and λ_{st} of the interfaces.

The upper insert in Fig. 5 shows the variation of the function $\Delta T_f(t)$ with λ_{st} for λ_{fs} = 300 W/(m^2 K). The dotted line separates the fast and slow parts of the response and demonstrates the optimal shape of the $\Delta T_f(t)$ dependence with a linear amplitude–frequency characteristic.

Note that λ_{st} on the order of (10^2–10^3) W/(m^2 K) is provided with appropriate adhesives. Therefore, the basic challenge in optimizing the film structures is to develop processes that provide a high thermal resistance at the film–substrate interface. To date, only few examples of obtaining λ_{fs} of desired order have been reported [5].

CONCLUSION

We presented the pulsed microwave technique for measuring the thermal kinetic coefficients in single-layer film structures. Unlike the other known methods of analyzing the bolometric response, it quantitatively traces both the temporal and the amplitude characteristics of the film temperature during heating. Its difference from the other methods of measuring the coefficients is that the film simultaneously serves as a heater and a thermometer.

Combined with a computer program, our technique is suitable for the express diagnostics of the thermal kinetic coefficients in film structures and also for monitoring the modifications of the transition region in the film under various process conditions. In addition, it can be useful in optimizing film bolometers in terms of power sensitivity and operating speed. Finally, it enables one to separate out direct thermal effects when electrodynamic phenomena in semiconductor and metal films and eddy dynamics in superconducting films are studied.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education and Science of Ukraine (project no. 2M/71-2000) and the Royal Academy of Sciences of Sweden.

REFERENCES

1. S. I. Khartsev, P. Johnsson, and A. M. Grishin, J. Appl. Phys. **87**, 2394 (2000).
2. A. Goyal, M. Rajeswari, R. Shreekala, *et al.*, Appl. Phys. Lett. **71**, 2535 (1997).
3. Z. M. Zang and A. Frenkel, J. Supercond. **7**, 871 (1994).
4. P. L. Richards, J. Appl. Phys. **76**, 1 (1994).
5. Ravi S. Prasher and Patrick E. Phelan, J. Supercond. **10**, 473 (1997).
6. A. V. Sergeev, A. D. Semenov, P. Kouminev, *et al.*, Phys. Rev. B **49**, 9091 (1994); A. Yu. Klokov, T. I. Galkina, and A. F. Plotnikov, Fiz. Tverd. Tela (St. Petersburg) **40**, 191 (1998) [Phys. Solid State **40**, 169 (1998)].
7. A. M. Grishin, Yu. V. Medvedev, and Yu. M. Nikolaenko, Fiz. Tverd. Tela (St. Petersburg) **41**, 1377 (1999) [Phys. Solid State **41**, 1260 (1999)].
8. C. Dekker, R. H. Koch, and A. Gupta, Physica C (Amsterdam) **185–189**, 1799 (1991).
9. Yu. D. Varlamov, A. N. Volkov, M. R. Predtechenskiĭ, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **5**, 1736 (1992).
10. H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon, Oxford, 1959; Nauka, Moscow, 1964).
11. P. C. Michael, J. U. Trefny, and B. Yarar, J. Appl. Phys. **72**, 107 (1992).
12. A. P. Ramírez, P. Schiffer, S. W. Cheong, *et al.*, Phys. Rev. Lett. **76**, 3188 (1996).
13. M. Castro, R. Burriel, and S. W. Cheong, J. Magn. Magn. Mater. **196–197**, 512 (1999).
14. I. G. Kozhevnikov and L. A. Novitskiĭ, *Thermophysical Properties of Materials under Low Temperatures* (Mashinostroenie, Moscow, 1982).
15. I. Z. San, D. W. Abraham, R. A. Rao, and C. B. Eom, Appl. Phys. Lett. **74**, 3017 (1999).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 121–124. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 125–128.

BRIEF COMMUNICATIONS

Generation of Electric Fields by Dust Grains during the Explosions of Condensed Explosive Charges in Air

M. A. Erokhin, A. A. Kondrat'ev, I. V. Litvinenko, and Yu. I. Matveenko

Zababakhin All-Russia Research Institute of Technical Physics, Russian Federal Nuclear Center, Snezhinsk, Chelyabinsk oblast, 456770 Russia

e-mail: yu.i.matveenko@vniitf.ru

Received September 19, 2000; in final form, May 10, 2001

Abstract—Results are presented from two-dimensional numerical simulations of the generation of electric signals during the explosions of chemical explosive charges in air. The calculated results are compared with the data from experiments carried out under the auspices of the International Science and Technology Center (project no. 835) at the Institute of Geosphere Dynamics (Russian Academy of Sciences), the Research Institute of Pulsed Technique (the Ministry of Atomic Energy of the Russian Federation), and the Zababakhin All-Russia Research Institute of Technical Physics, Russian Federal Nuclear Center (the Ministry of Atomic Energy of the Russian Federation). The electric field is assumed to be generated by charged dust grains produced during the explosion of a condensed chemical explosive charge. Numerical predictions based on the dusty plasma model for describing the parameters of dust grains are found to agree satisfactorily with experiment. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is well known that electric fields recorded during the explosions of chemical explosive (CE) charges are dipolar in character [1, 2]. Boronin *et al.* [3, 4] reviewed experimental and theoretical papers on this subject and proposed that the mechanism for generating such electric fields could be associated with explosively produced, negatively charged solid graphite grains. They also qualitatively explained the experimentally obtained time dependence of the strength of the electric fields generated during the explosions of CE charges with masses ranging from 10^{-3} to 5 kg.

The mechanism for generating electromagnetic pulses was investigated experimentally under the auspices of the International Science and Technology Center (project no. 835) at the Institute of Geosphere Dynamics (Russian Academy of Sciences), the Research Institute of Pulsed Technique (the Ministry of Atomic Energy of the Russian Federation), and the Zababakhin All-Russia Research Institute of Technical Physics, Russian Federal Nuclear Center (the Ministry of Atomic Energy of the Russian Federation). The experiments were carried out with spherical CE charges with a mass of 2.9 kg, made of 1 : 1 trinitrotoluol–hexogen alloy (TH-50/50). The charges were suspended by ropes at a height of 2 m above the Earth's surface and were detonated by flame ignition from above or at the center [5].

Our purpose here is to show that the results of numerical simulations based on the idea of generating charged dust grains during the explosion agree satisfactorily with the data obtained from measurements of the electric field component of electromagnetic waves in field experiments carried out at the Zababakhin All-Russia Research Institute of Technical Physics, Russian Federal Nuclear Center, in the summer of 1999.

THEORETICAL MODEL

The TH-50/50 CE has a negative oxygen balance; i.e., the explosion products (EPs) contain free carbon. During the explosion of the TH-50/50 CE, diamond dust and graphite soot are produced in the chemical reaction zone (at the chemical peak) within the detonation front over a time of about 1 μs [6, 7]. The kinetic processes of the production of seed diamond grains and of their subsequent growth finish ahead of the Chapman–Jouguet plane. In the same zone (i.e., at the chemical peak), the pressure is as high as hundreds of thousands of atmospheres, the temperature amounts to 0.35 eV, and the conductivity of the EPs, first, increases to $\sigma \approx 10$–$20\ \Omega^{-1}\ m^{-1}$ and then, after the passage of the detonation front, decreases by an order of magnitude during the time $t = 1$ μs.

Following [8, 9], we can estimate the grain charge q_p from the parameter values at the chemical peak in the region of the EPs. The grain charge turns out to be proportional to the grain potential ϕ_p with respect to the plasma: $q_p = eZ_p = C_p\phi_p$, where $C_p = 4\pi\varepsilon_0 R_p$ is the grain capacitance, R_p is the grain radius, and Z_p is the number of electrons collected by the grain (the grain charge number). The grains collect electrons and thus acquire

a negative charge until the electron heat flux onto the grain surface becomes equal to the ion heat flux, $Z_p e^2/(4\pi\varepsilon_0 R_p) = \beta k T_e$, where the quantity $\beta \approx 2$ depends logarithmically on the ion density, ion mass, and ion temperature and is approximately equal to two. We substitute the electron temperature T_e = 3800 K at the chemical peak into the last expression to obtain the grain charge number

$$Z_p = 4.2 \times 10^8 R_p. \quad (1)$$

The initial geometry of our numerical problem is as close as possible to the experimental one. A spherical CE charge with a mass of 2.9 kg is exploded at a height of 2 m above the Earth's surface. It is detonated either at the center (as is the case in the first experimental series) or at the top (as is the case in the second experimental series). The air pressure (or the EP pressure) is described by the gas-dynamic equations

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0, \quad \frac{\partial \mathbf{u}}{\partial t} = (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{\nabla p}{p},$$
$$\frac{\partial}{\partial t}\left(\rho\varepsilon + \frac{\rho u^2}{2}\right) + \nabla \cdot \left(\rho \mathbf{u}\left(\varepsilon + \frac{u^2}{2}\right) + p\mathbf{u}\right) = \rho Q, \quad (2)$$

where ρ, ε, and $\mathbf{u}$ are the density, internal energy, and velocity of the EPs (or air), and Q is the power released in an explosion.

The equations of state for air and for a TH-50/50 CE are presented in [10]. These equations and Eqs. (2) were solved together with the following equations of motion of the dust grains:

$$M_p \frac{d\mathbf{u}_p}{dt} = -6\pi\eta R_p(\mathbf{u}_p - \mathbf{u}), \quad (3)$$

where $\mathbf{u}_p$ is the velocity of a grain and $M_p = (4/3)\pi R_p^3 \rho_p$ and ρ_p = 2 g/cm^3 are its mass and density. In simulations, the dynamic viscosity was set to be constant, $\eta = 10^{-3}$ g/(cm s). Note that, in the equations of motion, the effect of the electric field on the grain dynamics is neglected, because it is insignificant in comparison with that of the forces of viscous friction of a grain against the EPs. The initial grain velocity was assumed to be zero.

Over the computation time (up to 5 s), the dust grains and EPs remained in a spherical region 4 m in diameter. On the other hand, the signals were recorded at a distance of 30 m from the explosion center. For these reasons, we calculated the electric signals in the dipole approximation, in which the electric field at the detectors has the form

$$\mathbf{E} = \frac{3\mathbf{R}(\mathbf{D} \cdot \mathbf{R})}{R^5} - \frac{\mathbf{D}}{R^3}.$$

Here, the dipole moment is equal to

$$\mathbf{D} = \mathbf{e}_z\left(\sum_k q_k z_k + \int \rho_q(r, z) z 2\pi r dr dz\right),$$

where R is the distance from the explosion center, q_k is the charge of a grain, z_k is its coordinate at the z axis, $\rho_q(r, z)$ is the space charge density in the region of the EPs, and (z, r) are cylindrical coordinates. In simulations, the dipole moment $\mathbf{D}$ was assumed to be axisymmetric, i.e., directed along the symmetry axis of the problem (the z-axis).

The gas motion was simulated in two-dimensional axisymmetric geometry (the symmetry axis is perpendicular to the Earth's surface and passes through the center of the spherical CE charge) using a particle-in-cell method on an adjustable nonuniform rectangular grid. The key elements of the simulation algorithm are as follows. (i) The particles are described by coordinates, velocity components, mass, internal energy, positive electric charge, and some other auxiliary quantities needed for an appropriate implementation of the code. (ii) An adjustable nonuniform rectangular grid that is used to model the particle motion in the computation region is constructed anew at the beginning of each time step. (iii) The grid is adjusted by dividing into four equal cells each spatial cell in which the number of particles turns out to be larger than a prescribed number. (iv) In each cell of the grid, the integral particle parameters, namely, the density ρ, mean velocity, specific internal energy ε, and pressure p, are calculated from the parameters of the particles within the cell. (v) The gas-dynamic equations (2) are integrated using the explicit difference scheme proposed by Harlow [11].

In simulations, the dust grains are produced and charged at the detonation front in the region of the EPs. In modeling the production and charging processes, the amount of dust grains originating in a cell with EPs is chosen to be proportional to the mass of the EP in the cell (the grains are assumed to be uniformly distributed over the cell volume) and the total grain charge in the cell is set equal in magnitude but opposite in sign to the charge of the EPs within it (so that the quasineutrality condition is satisfied at each time step).

Our approach to simulating the dynamics of charged grains assumes that they move in step with the EPs. At the beginning of each time step, the code determines the cell in which each of the grains is to occur; then, the velocity of the cell is used to solve the equations of motion (3).

NUMERICAL RESULTS AND COMPARISON WITH EXPERIMENT

The strength of the vertical component of the electric field was recorded at the Earth's surface at a distance of 30 m from the explosion center. In both series of experiments, the electric field strength was revealed

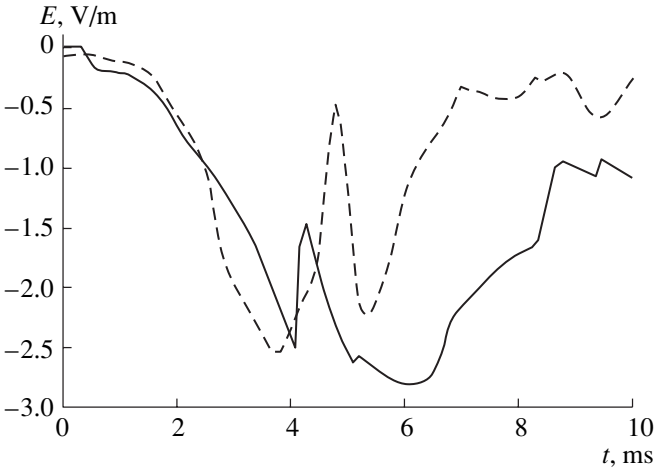


Fig. 1. Experimentally obtained (dashed curve) and numerically calculated (solid curve) time evolutions of the vertical component of the electric field during the explosion of a CE charge detonated at the center.

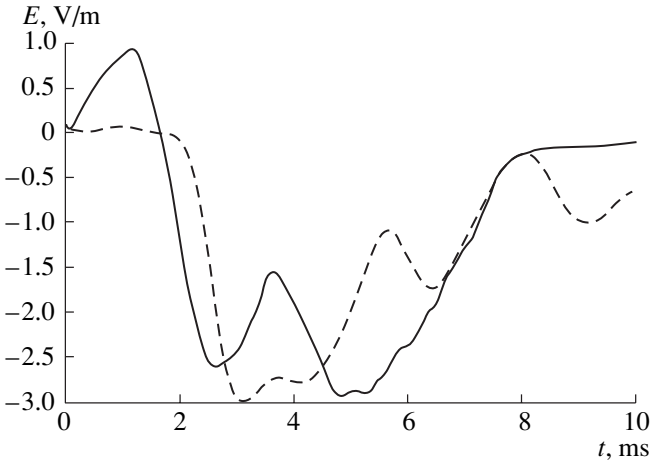


Fig. 2. Experimentally obtained (dashed curve) and numerically calculated (solid curve) time evolutions of the vertical component of the electric field during the explosion of a CE charge detonated at the top.

to have the same amplitude, E = 3 kV/m, but its time evolution was found to differ between the series (Figs. 1, 2).

Recall that we modeled the explosions of spherical CE charges detonated either at the center (first experimental series) or at the top (second experimental series). Qualitatively, the motion of EPs and dust grains can be described as follows. When a CE charge is detonated at a height of 2 m above the Earth's surface, the EPs expand to a radius of 100 cm during a time of 1 ms and then start to contract. By the time of 1 ms, a cloud of 9.5-μm-diameter dust grains expands to a radius of 20 cm. At a time of 1.5 ms, the shock wave (SW) reaches the Earth's surface. The SW reflected from the surface drives the EPs and dust grains into asymmetric motion, giving rise to the electric dipole moment and, accordingly, to the electric field. The main difference between the first and the second experimental series is that, in the latter case, the dipole moment is originally nonzero because the detonation is asymmetric.

The electric field strength depends on two parameters: the grain radius R_p and the total number N_p of dust grains. Note that the larger the grain radius, the longer the time t_f during which the electric signal rises to the peak amplitude. The reason for this is that the heavier grains are accelerated slower and stop moving later in comparison with the lighter grains. In simulations, we varied the grain radius from 0.1 to 50 μm in order to make the agreement as close as possible between the numerically calculated and experimentally obtained t_f values. The grain charge (1) is uniquely determined by the grain radius. We also varied the total number N_p of dust grains in the EP in order to achieve agreement between the calculated and measured amplitudes of the electric field strength.

The results of numerical calculations of the electric field strength are illustrated in Figs. 1 and 2. A satisfactory agreement between the calculated and measured electric fields is achieved for the grain radius R_p = 9.5 μm and the total grain charge $q_{tot} = 3.3 \times 10^{-5}$ C. At the time at which the electric dipole moment of the system of EPs and dust grains reaches its maximum value, $d_{max} = 9 \times 10^{-7}$ C m, the effective distance L by which the charges are separated is equal to $L = d_{max}/q_{tot}$ = 2.7 cm.

Substituting the grain radius R_p = 9.5 μm into expression (1) yields the amount of electrons collected by a grain, $Z_p = 4.3 \times 10^3$. In this case, the total amount of dust grains is $N_p = 4.8 \times 10^{10}$ and the total dust mass is 340 g, which constitutes 11.7% of the total mass of the EPs. This mass value agrees with the yield of diamond dust (up to 10% [6, 7]) in special-purpose explosive devices.

CONCLUSION

We have investigated the mechanism by which EPs acquire an electric dipole moment during the explosions of TH-50/50 CE charges above the Earth's surface. Our two-dimensional numerical simulations show that the spatiotemporal behavior of the electric signal can be satisfactorily described by choosing the value R_p = 9.5 μm for the grain radius and by assuming that the total mass of the dust grains is 11.7% of the total mass of the EPs.

ACKNOWLEDGMENTS

We are grateful to our colleagues from the Institute of Geosphere Dynamics (Russian Academy of Sciences) and the Research Institute of Pulsed Technique

(the Ministry of Atomic Energy of the Russian Federation) for providing us with the experimental data. We also thank M.M. Gorshkov and V.N. Pristavko for useful discussions of the results obtained. This work was supported in part by the International Science and Technology Center, project no. 835.

REFERENCES

1. W. H. Andersen and C. L. Long, J. Appl. Phys. **36**, 1494 (1965).
2. L. M. Gorshunov, G. P. Konopenko, and E. I. Sirotin, Zh. Éksp. Teor. Fiz. **53**, 818 (1967) [Sov. Phys. JETP **26**, 500 (1968)].
3. A. P. Boronin, V. N. Kapinos, S. A. Krenev, and V. N. Mineev, Fiz. Goreniya Vzryva **26**, 110 (1990).
4. A. P. Boronin, V. N. Kapinos, and S. A. Krenev, Fiz. Goreniya Vzryva **26**, 117 (1990).
5. *Proceedings of Russian-American Workshop "Electromagnetic Pulses from Chemical Explosions" by MNTTs Project no. 835-98* (NIIIT, Moscow, 1999), No. O-1575.
6. N. R. Geiner, D. S. Phillips, J. D. Jonson, and F. Volk, Nature **333**, 440 (1988).
7. A. I. Lyamkin, E. V. Petrov, A. P. Ershov, *et al.*, Dokl. Akad. Nauk SSSR **302**, 611 (1988) [Sov. Phys. Dokl. **33**, 705 (1988)].
8. V. N. Tsytovich and U. Angelis, Phys. Plasmas **6**, 1093 (1999).
9. V. N. Tsitovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys. Usp. **40**, 53 (1997)].
10. V. F. Kuropatenko, Fiz. Goreniya Vzryva **26** (6), 112 (1989).
11. F. H. Harlow, in *Fundamental Methods in Hydrodynamics*, Ed. by B. Alder, S. Fernbach, and M. Rotenberg (Academic, New York, 1964; Mir, Moscow, 1967).

Translated by G. Shepekina

Technical Physics, Vol. 47, No. 1, 2002, pp. 125–128. Translated from Zhurnal Tekhnichesko ĭ Fiziki, Vol. 72, No. 1, 2002, pp. 129–132.

BRIEF COMMUNICATIONS

The Permeability and Remanent Magnetization of a Randomly Inhomogeneous Two-Phase Medium

V. V. Bakaev*, A. A. Snarskiĭ, and M. V. Shamonin***

** H. ROSEN Engineering GmbH, Am Seitenkanal 8, D-49811 Lingen (Ems), Germany*

*** National Technical University of Ukraine, Kiev Polytechnical Institute, ul. Peremogi 37, Kiev, 03056 Ukraine*

e-mail: asnar@phys.carrier.kiev.ua

Received October 30, 2000

Abstract—A randomly inhomogeneous composite consisting of two, ferromagnetic and nonmagnetic (para- or diamagnetic), phases is considered. The dependence of the effective permeability of the composite on the concentration of the ferromagnetic phase and on the applied magnetic field is found for the case of the negligible hysteresis loop. When the hysteresis loop is appreciable, the remanent magnetization as a function of the ferromagnet concentration is calculated. © *2002 MAIK "Nauka/Interperiodica"*.

Magnetic composites are attracting widespread attention because of their intriguing properties. For example, the size of the hysteresis loop may vary from negligibly small to rather large. The composites may form the basis for creating new macroscopically inhomogeneous materials whose properties can be controlled in wide ranges (see, e.g., [1, 2]). Of great interest are also the giant Hall effect and the effect of giant magnetoresistance ([3, 4] and Refs. therein).

The basic characteristics of inhomogeneous media are effective coefficients, in particular, the effective permeability μ^e for magnetic composites. By definition, μ^e is the coefficient that relates the volume-averaged magnetic field strength H and induction B:

$$\langle \mathbf{B} \rangle = \mu^e(\langle \mathbf{H} \rangle)\langle \mathbf{H} \rangle, \qquad (1)$$

where $\langle \ldots \rangle = V^{-1}\int \ldots dV$ and V is the volume of averaging.

For random-structure composites, the effective coefficients are very difficult to calculate and the problem does not have the general solution [5]. There are only few approximations that make it possible to derive analytical dependences of the effective coefficients on the phase arrangement, degree of inhomogeneity, etc. For a low concentration of the inclusions, when their interactions can be neglected, the Maxwell approximation is valid. In real composites, however, this approximation fails as a rule. The best approximation among those taking into account the interaction between the inclusions is the Bruggeman–Landauer (BL) self-consistent field approximation [6–8]. It gives the reasonable values of all parameters everywhere except in the immediate vicinity of the percolation threshold at high degrees of inhomogeneity. In this case, the percolation theory must be applied [8, 9].

In this work, we consider a 3D randomly inhomogeneous medium where one phase is ferromagnetic and the other is nonmagnetic (dia- or paramagnetic phase whose susceptibility is negligibly small). Although the problem of determining the effective permeability μ^e [see (1)] is formally similar to the problems of determining the effective electric conductivity ($\langle \mathbf{j} \rangle = \sigma_e \langle \mathbf{E} \rangle$, where $\mathbf{j}$ is the current density and $\mathbf{E}$ is the field; div$\mathbf{j} = 0$, curl $\mathbf{E} = 0$) and effective thermal conductivity ($\langle \mathbf{q} \rangle = \kappa_e \langle -\nabla T \rangle$, where $\mathbf{q}$ is the thermal flux density and ΔT is the temperature gradient; div $\mathbf{q} = 0$), it differs from them at least in two fundamental aspects, which greatly complicate the analysis. First, the local effective permeability $\mu(H)$ features specific nonlinearity (see insert in Fig. 1), which is neither weak (such as Ohm's law under low fields) nor strong (such as the power j–E characteristic) [10, 11]. At high magnetic fields, the relative permeability is close to unity; that is, a medium consisting of ferromagnetic and nonmagnetic phases is weakly inhomogeneous. Under moderate fields, however, the relative permeability may be very high (500 or more) and the medium thus becomes strongly inhomogeneous. Second, if the remanent magnetization in the ferromagnetic phase is appreciable, the notion of permeability becomes ambiguous.

We will consider first the case when the hysteresis loop is negligibly small, that is, when the function $\mu_1 = \mu_1(H)$ in the ferromagnetic phase is uniquely defined. Then, the problem of determining the remanent magnetization will be treated.

To find the dependence μ^e on the external field, we will employ the BL approximation modified [12, 13] for the case of the power nonlinear j–E characteristic, $j = \chi E^\beta$. Following [12, 13] and taking into account that the local field $\mathbf{H}(\mathbf{r})$ inside the inclusions is constant [14], we will characterize the nonlinear phase with

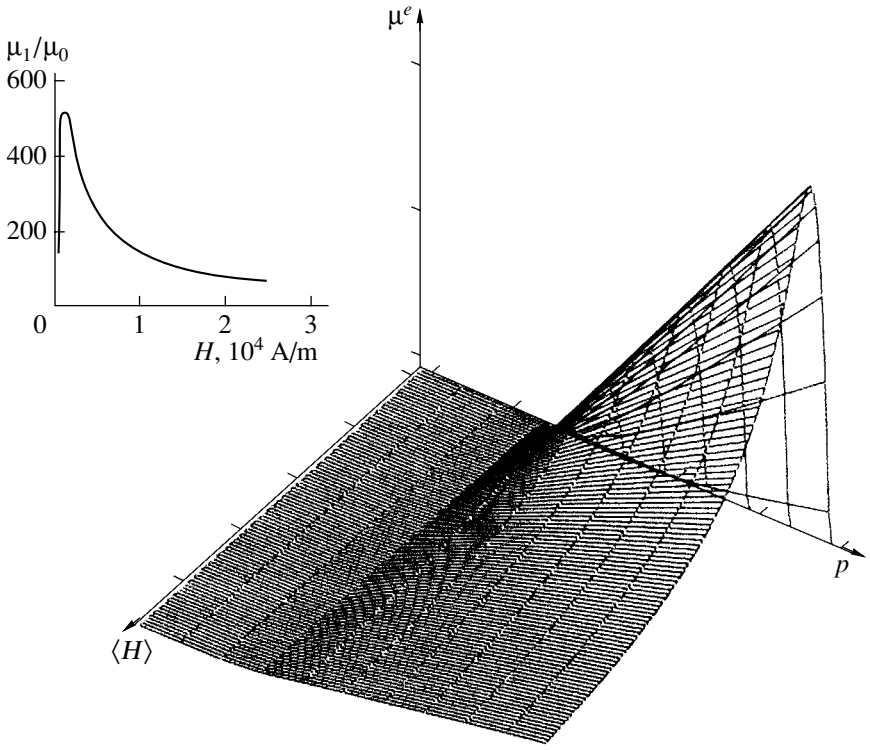


Fig. 1. Effective permeability vs. concentration of the ferromagnetic phase p and average field strength $\langle H \rangle$. The insert shows the field dependence of the relative permeability of the ferromagnetic phase.

$\mu_1(H)$ by some constant permeability $\tilde{\mu}_1$:

$$\tilde{\mu}_1 = \langle \mu_1(H) \rangle_1, \tag{2}$$

where averaging is over the volume.

Replacement (2) means that the composite now consists of two "linear" phases with $\tilde{\mu}_1$ and μ_2 (the latter practically equals μ_0). Thus, μ^e can be found with the conventional BL approximation [6–8]

$$\mu_\perp^e = \frac{1}{4}\{(3p-1)\tilde{\mu}_1 + (2-3p)\mu_2 + \sqrt{[(3p-1)\tilde{\mu}_1 + (2-3p)\mu_2]^2 + 8\mu_2\tilde{\mu}_1}\}, \tag{3}$$

where p is the concentration of the ferromagnetic phase.

To find μ^e as a function of $\langle H \rangle$, we take advantage of the relationship [12, 13]

$$\langle H^2 \rangle_1 = \frac{\langle H^2 \rangle}{p}\frac{\partial \mu^e}{\partial \tilde{\mu}_1}, \tag{4}$$

which is easily found from the equality [15]

$$\langle \mathbf{B} \cdot \mathbf{H} \rangle = \langle \mathbf{B} \rangle \cdot \langle \mathbf{H} \rangle \tag{5}$$

and definition (1) of the effective coefficient (provided that the size of the sample is larger than the dimension of averaging).

It is also necessary to express $\tilde{\mu}_1$ through $\langle H^2 \rangle_1$. To do this, we put

$$\tilde{\mu}_1 = \langle \mu(H) \rangle_1 \approx \mu_1(\sqrt{\langle H^2 \rangle_1}). \tag{6}$$

Such an approximation has proved to be adequate in calculating the electrical conductivity from the power j–E characteristics of the phases [12, 13]. Substituting (6) into (3) and (3) into (4) yields the nonlinear equation for $\langle H^2 \rangle_1$:

$$\langle H^2 \rangle_1 = \frac{\langle H^2 \rangle}{4p}\left[3p - 1 + \frac{[(3p-1)\tilde{\mu}_1 + (2-3p)\mu_2](2-3p) + 4\mu_2}{\sqrt{[(3p-1)\tilde{\mu}_1 + (2-3p)\mu_2]^2 + 8\tilde{\mu}_1\mu_2}}\right]. \tag{7}$$

Having found the field $\langle H^2 \rangle_1$ from (7) and having substituted it into (3), we find the effective permeability μ^e of the ferromagnetic composite as a function of the concentration, applied magnetic field $\langle H \rangle$, and nonlinearity parameters.

By way of example, let us analyze the local permeability curve $\mu_1 = \mu_1(H)$ depicted in the insert to Fig. 1. This figure shows the dependence $\mu^e = \mu^e(p, \langle H \rangle)$ obtained from (7) and (8). At $p = 1$, the field dependence of μ^e coincides with the field dependence of the

permeability of the purely ferromagnetic phase, as it must. For fields at which $\mu_1 = \mu_1(H) \gg 1$, the concentration dependence of μ^e sharply grows in going over the percolation threshold, which is $p_c \approx 0.2$ for the 3D case.

Now let us consider the case when the ferromagnetic phase exhibits an appreciable hysteresis loop. Here, the notions of local and effective permeabilities become ambiguous. Bearing in mind that large-hysteresis ferromagnetics are used largely as permanent magnets, we will look for the dependence of the remanent magnetization B_r of the composite on the concentration of the ferromagnetic phase and consider the upper (saturating) part of the loop $B_m = B_m(H)$ (Fig. 2). It is assumed that, upon magnetization, the applied field $\langle H \rangle$ first reaches values equal to or higher than H_s (the field of technical saturation [16]) and then drops to zero.

In our approximate approach to calculating the effective parameters of the composite, all inclusions of a given phase are in the "same position." The interaction between them is included in the self-consistency scheme, but the local fields inside the inclusions of a given phase are the same. Then, the local remanent magnetization b_r in all of the ferromagnetic inclusions is also the same. Since we are now interested in B_r (i.e., in the range of low magnetic fields, $H \approx 0$), the saturating magnetization curve can be approximated by the linear dependence (Fig. 2)

$$B(H) = \mu_d H + b_r, \tag{8}$$

where $\mu_d = dB(0)/dH$ is the differential zero-field permeability for the saturating part of the magnetization curve.

Thus, the dependence $B = B(H)$ for the ferromagnetic phase has the form of (8) and for the nonmagnetic phase, $B(H) = \mu_0 H$ as before. The problem of finding the effective coefficients for such local laws is thus reduced to that of finding the effective thermopower (up to designations). In fact, in the presence of thermoelectricity,

$$\mathbf{j} = \sigma\mathbf{E} + \sigma\alpha(-\nabla T), \tag{9}$$

where $\mathbf{j}$ is the current density, $\mathbf{E}$ is the electric field strength, σ is the electrical conductivity, α is the thermopower, and ∇T is the temperature gradient.

Consider the case when (i) the medium is uniform in terms of the thermal conductivity (the thermal conductivities of the phases κ_1 and κ_2 equal each other) and (ii) $\alpha_2 = 0$ (the thermal emf in the second phase is absent). Conditions (i) and (ii) mean that the gradient $\nabla T = \text{const}$ (let it equal g). Let the thermopower of the first phase be

$$\alpha_1 = -b_r/\sigma_1 g, \tag{10}$$

then,

$$\begin{aligned} j &= \sigma_1 E + b_r \text{ (first phase)}, \quad B = \mu_d H + b_r, \\ j &= \sigma_2 E \text{ (second phase)}, \quad B = \mu_0 H. \end{aligned} \tag{11}$$

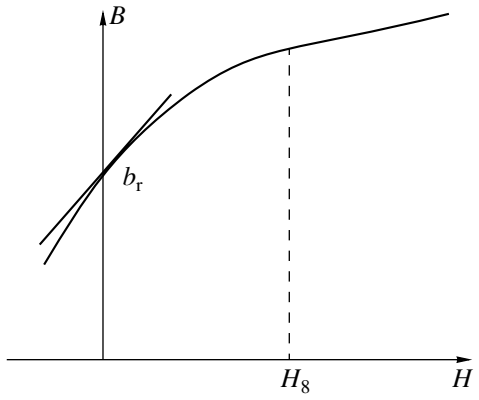


Fig. 2. The upper part of the hysteresis loop. b_r, remanent magnetization; H_s, field of technical saturation.

It is seen that (11) coincides with the statement of the problem of remanent magnetization up to designations ($j \rightleftharpoons B$, $E \rightleftharpoons H$, $\sigma_i \rightleftharpoons \mu_i$) [see (8) and below].

Thus, to find the remanent magnetization of the composite, it is necessary to determine the effective thermopower α_e (in terms of the thermoelectric problem); then, B_r will equal $\alpha_e(\alpha_1 = -b_r/\sigma_1 g)\sigma_e g$ up to designations. The expression for α_e in the BL approximation has been obtained in [17, 18]; in our case, it can be written as

$$\alpha_e = p\alpha_1\sigma_1 \frac{2\sigma_e + \sigma_2}{2p(\sigma_1 - \sigma_2)\sigma_e + \sigma_1\sigma_2}, \tag{12}$$

where the effective electrical conductivity is also given in terms of the SCFMA approximation:

$$\begin{aligned} \sigma_e = \frac{1}{4}[&(3p-1)\sigma_1 + (2-3p)\sigma_2 \\ &+ \sqrt{[(3p-1)\sigma_1 + (2-3p)\sigma_2] + 8\sigma_1\sigma_2}. \end{aligned} \tag{13}$$

In view of the aforesaid and making the replacements $\sigma_1 \rightleftharpoons \mu_d$, $\sigma_2 \rightleftharpoons \mu_0$, and $\sigma_e \rightleftharpoons \mu^e$, we can write for the remanent magnetization

$$B_r = b_r p \frac{\mu^e}{\mu_0} \frac{2\dfrac{\mu^e}{\mu_0} + 1}{2\left[p\left(\dfrac{\mu_d}{\mu_0} - 1\right) + 1\right]\dfrac{\mu^e}{\mu_0} + \dfrac{\mu_d}{\mu_0}}. \tag{14}$$

Figure 3 shows the dependence of the remanent magnetization on the concentration of the ferromagnetic phases.

It has been experimentally found (Figs. 14–16 in [1]) that μ^e sharply grows with p in some range of the concentrations, implying the presence of the percolation transition. According to [1], the percolation threshold is near $p_c = 0.25$. It is known that the BL approximation gives incorrect values of the critical indices of the effective coefficients but qualitatively predicts an

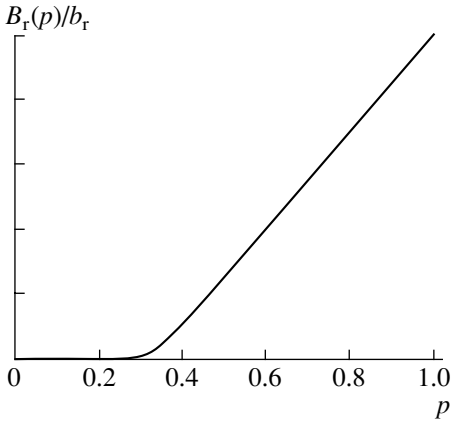


Fig. 3. Remanent magnetization B_r vs. concentration of the ferromagnetic phase.

increase in the effective coefficients when the concentration exceeds the percolation threshold. The nonlinear BL approximation also predicts a sharp growth of μ^e at the same concentrations (Fig. 1). However, the quantitative parameters of this increase, as well as the run of the curve near the threshold, depend on the applied magnetic field (Fig. 1). In terms of the percolation theory, this means that the critical indices depend on the magnetic field. However, in [1], neither the field dependences of the permeability of the ferromagnetic phase nor the fields at which the concentration dependence of μ^e were recorded were reported. Therefore, we cannot compare our analytical data with the experiment. Note that the remanent magnetization also behaves in a percolation-like manner, as follows from Fig. 3. Finally, of interest is the unevident fact that the field at which μ^e reaches a maximum depends on the concentration of the ferromagnetic phase.

ACKNOWLEDGMENTS

The authors are indebted to P. Rosen, T. Beuker, and J. Rosenmoeller for the fruitful discussion and Pak-Ming Hui and D. Kalyon for sending the preprints.

V. Bakaev and A. Snarskiĭ thank H. ROSEN Engineering GmbH for the hospitality during the preparation of this article.

REFERENCES

1. T. J. Fiske, H. S. Gokturk, and D. M. Kalyon, J. Mater. Sci. **32**, 5551 (1997).
2. H. S. Gokturk, T. J. Fiske, and D. M. Kalyon, J. Appl. Phys. **73**, 5598 (1993).
3. Physica A (Amsterdam) **241** (1–2) (1997); Physica B (Amsterdam) **279** (1–3) (1999).
4. S. K. Wong, B. Zhao, T. K. Ng, *et al.*, Eur. J. Phys. B **10**, 481 (1999).
5. K. J. Binns, P. S. Lawrenson, and C. W. Trowbridge, *The Analytical and Numerical Solution of Electric and Magnetic Fields* (Wiley, Chichester, 1992).
6. R. Landauer, J. Appl. Phys. **23**, 779 (1952).
7. D. A. Bruggeman, Ann. Phys. (Leipzig) **24**, 636 (1935).
8. J. P. Clerc, G. Giraud, J. M. Laugier, and J. M. Luck, Adv. Phys. **39**, 191 (1990).
9. D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1992), p. 181.
10. S. W. Kenkel and J. P. Straley, Phys. Rev. Lett. **49**, 767 (1982).
11. J. P. Straley and S. W. Kenkel, Phys. Rev. B **29**, 6299 (1984).
12. P. M. Hui, P. Cheung, and Y. R. Kwong, Physica A (Amsterdam) **241**, 301 (1997).
13. P. M. Hui, Y. F. Woo, and W. M. V. Wan, J. Phys. C **7**, L593 (1995).
14. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1960), p. 620.
15. A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 110 (1970) [Sov. Phys. JETP **32**, 63 (1971)].
16. D. Jiles, *Introduction to Magnetism and Magnetic Materials* (Chapman and Hall, London, 1991), p. 440.
17. I. Webman, J. Jortner, and M. H. Cohen, Phys. Rev. B **16**, 2959 (1977).
18. B. Ya. Balagurov, Fiz. Tekh. Poluprovodn. (Leningrad) **16**, 1870 (1982) [Sov. Phys. Semicond. **16**, 1204 (1982)].

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 129–131. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 133–135.

BRIEF COMMUNICATIONS

Characteristics of the Reconstructed Image upon Multiplexing On-Axis Holograms

N. M. Ganzherli, I. A. Maurer, and D. F. Chernykh

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

Received April 16, 2001

Abstract—A comparative analysis of the experimental data on recording and multiplexing holograms, including speckle holograms, by an axial scheme is presented. The contrast and signal-to-noise ratio in the images reconstructed are compared. It is demonstrated that the thickness of the photosensitive material has an effect on the image quality during the recording of filamentary holograms. *© 2002 MAIK "Nauka/Interperiodica".*

Recently, we have reported experimental data on recording and multiplexing on-axis holograms, including shift speckle holograms [1–4]. In this paper, we present the results of experiments aimed at estimating the quality of the reconstructed images upon on-axis multiplexing. The quality was estimated by the image contrast and the signal-to-noise ratio. For recording and multiplexing the on-axis holograms, we used the experimental setups shown in Figs. 1a and 1b. The latter variant was used for recording shift speckle holograms.

In recording the holograms by the axial scheme (Fig. 1a), objects were transparencies representing a number of equiwide concentric transparent rings on a nontransparent screen. The entire system had a fixed diameter and was surrounded by a transparent ring providing the reference beam.

The method for recording on-axis holograms (Fig. 1a) was proposed by Denisyuk for recording and multiplexing holograms in 3D media [1]. Owing to the specific distribution of light near the focal plane of the lens, such holograms form as thin filaments running along the optical axis and being symmetric about this axis and with respect to the focal plane. For each of the transparency rings, the length of the filament is proportional to the wavelength and to the square of the ratio between the focal length and the ring radius. Therefore, the thickness of the photosensitive medium is an important parameter in recording on-axis holograms.

The recording density upon multiplexing depends primarily on the holograms dimensions. In the case considered, the hologram recorded in the focal plane is formed by the interference of the Fourier spectrum of the outer (reference) ring with that of the object, which is a set of transparent concentric rings. The dimensions of the holograms both in the focal plane and along the optical axis depend on the sizes of the overlapping spectral areas, where the intensity of light is sufficient for recording at a given exposure.

Consider the Fourier spectrum of a specific transparency in the focal plane of the lens L_2 (Fig. 1a). The object has six transparent rings with the width 1 mm and the inner diameters 45, 40, 35, 30, 25, and 20 mm and is surrounded by the reference ring of the same width and the inner diameter 55 mm. Figures 2a and 2c show the computed radial distributions of the intensity in the Fourier spectra of the reference ring and the object, respectively. It is seen that the spectral intervals coincide but the intensity of the reference signal is five times smaller. In the calculations, the intensity of the

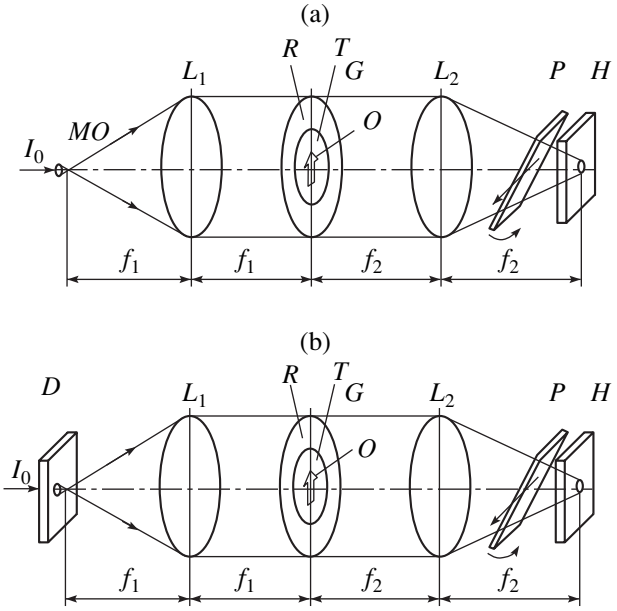


Fig. 1. Axial schemes for recording and reconstructing (a) 3D and (b) shift speckle holograms. I_0, laser beam; *MO*, microobjective; *D*, diffuser; L_1 and L_2, lenses; *T*, transparency with object image *O*; *G*, chopper, which shuts light transmitted through transparency *T* during reconstruction; *R*, reference ring; *P*, rotating plane-parallel plate; and *H*, photosensitive material.

1063-7842/02/4701-0129$22.00 © 2002 MAIK "Nauka/Interperiodica"

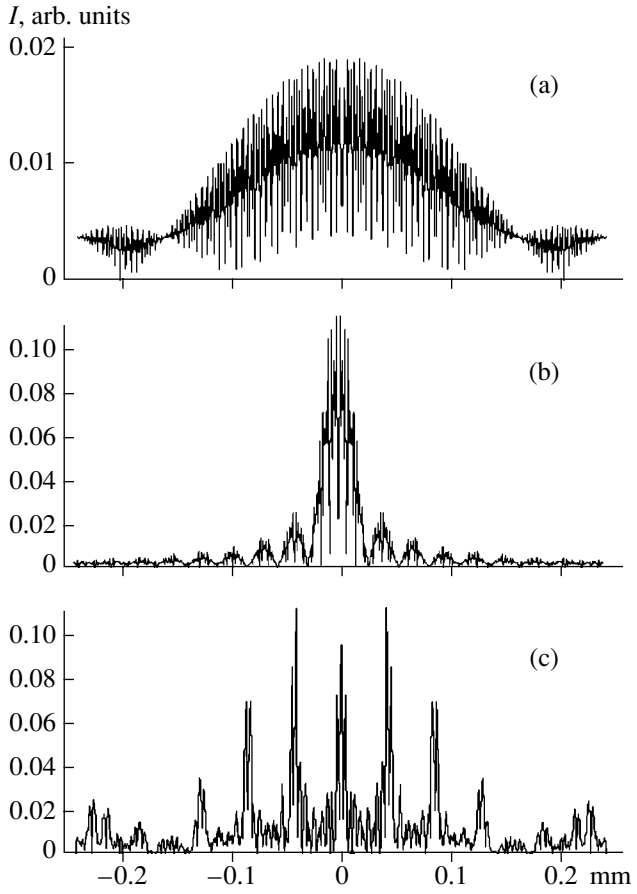


Fig. 2. Fourier spectra of the reference ring with the width (a) 1 and (b) 10 mm. (c) Fourier spectrum of the object consisting of six equiwide rings.

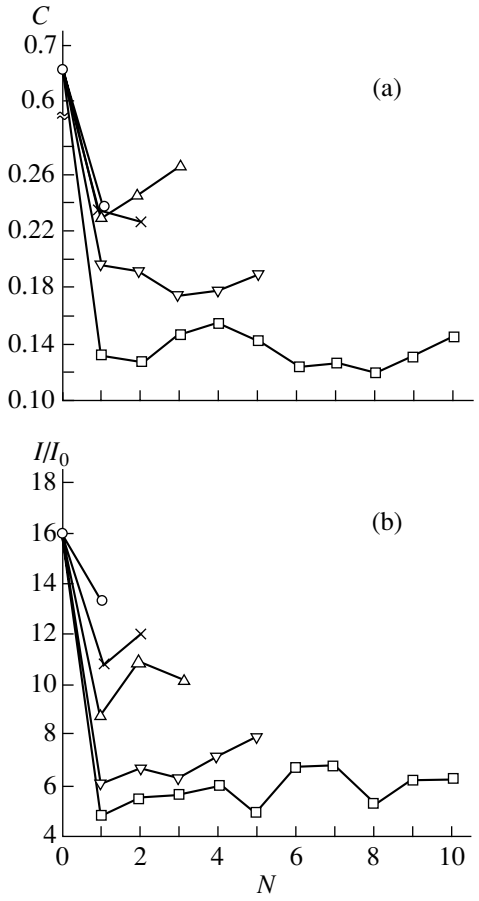


Fig. 3. (a) Contrast and (b) signal-to-noise ratio of the multiplexed images vs. the order of multiplexing.

incident beam was assumed to be equal to unity. To attain the maximum contrast of the interference pattern, one must equalize the intensities of the spectra. An increase in the width of the reference ring raises the reference beam intensity but narrows the central maximum of the spectrum (Fig. 2b). Hence, the overlap strinks. In the experiments, we attenuated the object beam intensity by a filter.

As a recording medium, we used PFG-03M holographic photoplates, which must be photochemically processed after exposure. We also employed real-time media, which do not require postexposure processing: lithium niobate crystals and self-developing glycerol-containing dichromated gelatin layers [5].

The images reconstructed were detected by a 795×596-pixel CCD array and transferred to a computer for processing. No preliminary preprocessing of the images to improve the contrast and suppress the noise was accomplished.

Comparative experiments on recording the same object in the media of various thickness revealed that the quality of the image reconstructed is improved with increasing thickness of the medium. The images with the highest contrast (0.48) and the lowest noise (a mean signal-to-noise ratio of 8.3) were obtained in the 1.31-mm-thick lithium noibate crystals. The holograms recorded in a 240-μm-thick layer of self-developing glycerol-containing dichromated gelatin exhibited the lower contrast (0.38) and the higher noise level (a signal-to-noise ratio of 6.3). The highest noise level (a signal-to-noise ratio of 5.3) and the lowest contrast (0.25) were typical of the holograms recorded in the PFG-03M photoplates with the thickness of the emulsion layer 7 μm. However, the quality of the image reconstructed by the photoplates was reasonable over the entire image area. Conversely, the holograms recorded in the lithium niobate and the gelatin layer had the nonuniform image quality, which is related to the volume inhomogeneity of the material.

In the experiments with the PFG-03M photoplates, we used the objects with a varying number (from 3 to 11) of transparent concentric rings of the same width. The increase in the number of the rings with the total size of the object remaining unchanged worsens the resolution of the rings in the image reconstructed, lowers the contrast, and raises the noise.

In multiplexing shift speckle holograms with the scheme shown in Fig. 1b, the highest order of multiplexing was obtained for the holograms recorded in gelatin layer. With a shift of the photosensitive material by several microns with the use of a plane-parallel plate P introduced into the object and reference beams, we recorded the holograms of each successive object. Real-time recording made it possible to visually determine the exposure time for each successive hologram. The total number of the holograms recorded ranged from 12 to 15, depending on the layer thickness and the exposure of the first hologram. Each successive image reconstructed had a higher noise level and a lower contrast.

We managed to record only ten holograms on the photoplates. The disadvantage of this material is the need for photochemically fixing the records, which is accompanied by the shrinkage of the emulsion and adversely affects the image reconstructed. In addition, it is difficult to determine the exposure time for each successive hologram and, accordingly, to use the entire dynamic range of the photomaterial. Figures 3a and 3b show, respectively, the contrast and the signal-to-noise ratio for the images reconstructed upon multiplexing the shift speckle holograms vs. the order of multiplexing N varying from 1 to 10.

ACKNOWLEDGMENTS

We thank P.V. Granskiĭ for the assistance in computer processing of the reconstructed images and M.V. Monakhova for preparing the gelatin layers.

This work was supported by the Russian Foundation for Basic Research, grant nos. 00-15-96771 and 99-02-18481.

REFERENCES

1. Yu. N. Denisyuk, N. M. Ganzherli, I. A. Maurer, and S. A. Pisarevskaya, Opt. Spektrosk. **86**, 1023 (1999) [Opt. Spectrosc. **86**, 922 (1999)].
2. Yu. N. Denisyuk, N. M. Ganzherli, I. A. Maurer, and S. A. Pisarevskaya, in *Proceedings of the Second Baikal School on Fundamental Physics, Irkutsk, 1999*, Vol. 1, p. 55.
3. Yu. N. Denisyuk, N. M. Ganzherli, S. P. Konop, *et al.*, Proc. SPIE **3956**, 326 (2000).
4. Yu. N. Denisyuk, N. M. Ganzherli, and D. F. Chernykh, Pis'ma Zh. Tekh. Fiz. **26** (9), 25 (2000) [Tech. Phys. Lett. **26**, 369 (2000)].
5. 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 A. Chikishev

Technical Physics, Vol. 47, No. 1, 2002, pp. 13–17. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 15–19.

On Interaction of Two Closely Spaced Charged Conducting Spheres

E. A. Shcherba, A. I. Grigor'ev, and V. A. Koromyslov

Demidov State University, Sovetskaya ul. 14, Yaroslavl, 150000 Russia

e-mail: grig@uniyar.ac.ru

Received March 7, 2001; in final form, May 10, 2001

Abstract—Electrostatic interaction between two charged conducting spheres is analyzed in the case of a small spacing between them, when the polarization effects are significant. It is shown that short-range polarization forces result in the attraction of the like-charged spheres. At a sufficiently small spacing, this attraction replaces repulsive forces acting on like charges. © *2002 MAIK "Nauka/Interperiodica".*

1. The problem of electrostatic interaction between two closely spaced charged particles (drops) in an electric field is a key issue in calculating the coagulation parameters in charged aerodisperse systems of natural and artificial origin, as well as in deducing disintegration law for drops unstable against self-charge and polarization charges [1–8]. Generally, the electrostatic interaction of two both closely and widely spaced charged particles is treated in terms of the Coulomb law. On the other hand, it is well known [9] that polarization interaction becomes of paramount importance when the particles are closely spaced (the situation typical of the coagulation and disintegration of drops). Eventually, like-charged particles are attracted to, rather than repelled from, each other for certain relations between their charges and sizes. However, this effect has been neglected, since the analytical expressions describing the force and energy of electrostatic interaction between charged particles in view of polarization effects involve awkward series inconvenient for practical calculations. However, the recently developed software suites for analytical computing have made possible qualitative and quantitative analysis of electrostatic interaction between closely spaced charged spherical particles. Also, the procedure of simplifying the poorly convergent series in the exact analytical expression for the energy of interaction between two charged conducting spheres [9] has been proposed [10].

2. Consider two conducting spheres of radii R_1 and R_2 and charges q_1 and q_2. Their center distance is l. Due to the electrostatic induction, the charge of the first sphere induces the electrostatic image Q_{11} in the other. This image, in its turn, induces the image q_{12} in the first sphere. Thus, the field strength at an arbitrary point A between the spheres on the axis connecting their centers is produced by the infinite number of the image charges on both spheres. If only the first sphere is charged ($q_2 = 0$), the field strength at the point A is given by [9]

$$E_{A_1} = k\left(\sum_n^\infty \frac{q_{1n}}{r_{1n}^2} + \sum_n^\infty \frac{Q_{1n}}{R_{1n}^2}\right), \quad k \equiv \frac{1}{4\pi\varepsilon_0},$$

where

$$q_{1n} = q_1 \frac{\gamma \sinh\beta}{\sinh(n\beta)[\gamma + \sinh[(n-1)\beta]/\sinh(n\beta)]},$$

$$Q_{1n} = -q_n \frac{\gamma \sinh\beta}{r(1+\gamma)\sinh(n\beta)},$$

$$r_{1n} = \frac{1}{1+\gamma} - r + \frac{r\gamma}{\gamma + \sinh[(n-1)\beta]/\sinh(n\beta)},$$

$$R_{1n} = r - \frac{1}{1+\gamma} - \frac{r\gamma[\gamma + \sinh[(n-1)\beta]/\sinh(n\beta)]}{1+\gamma^2+2\gamma\cosh\beta},$$

$$r \equiv \frac{l}{R_1 + R_2}, \quad \gamma \equiv R_2/R_1.$$

Here, the parameter β is related to the center distance of the spheres by the formula

$$\cosh\beta = \frac{r^2(1+\gamma)^2 - (1+\gamma^2)}{2\gamma}.$$

Now consider the case when the first sphere is neutral ($q_1 = 0$) and the other has the charge q_2. Then, for the field strength at the point A, we have

$$E_{A_2} = k\left(-\sum_n^\infty \frac{q_{2n}}{r_{2n}^2} + \sum_n^\infty \frac{Q_{2n}}{R_{2n}^2}\right),$$

$$q_{2n} = q_2 \frac{\sinh\beta}{\sinh\beta[1 + \gamma\sinh[(n-1)\beta]/\sinh(n\beta)]},$$

$$Q_{2n} = -q_2\frac{\sinh\beta}{r(1+\gamma)\sinh(n\beta)},$$

$$r_{2n} = -\frac{1}{1+\gamma} + \frac{r}{1+\gamma\sinh[(n-1)\beta]/\sinh(n\beta)},$$

$$R_{2n} = \frac{1}{1+\gamma} - \frac{r[1+\gamma\sinh[(n-1)\beta]/\sinh(n\beta)]}{1+\gamma^2+2\gamma\cosh\beta}.$$

If the spheres have different charges, the field strength at A is

$$E_A = E_{A_1} + E_{A_2}.$$

The potential energy of interaction between the spheres, as well as the force acting on either of them, can be represented in the form [9, 10]

$$W = \frac{1+\gamma}{2\alpha}\frac{\alpha^2 c_{11} - 2\alpha c_{12} + c_{22}}{c_{11}c_{22} - c_{12}^2}, \quad \alpha \equiv q_2/q_1, \qquad (1)$$

$$F_l = -\frac{\partial W}{\partial l}, \qquad (2)$$

where c_{ij} are the capacitive coefficients. In units of R_1/k, they are expressed as

$$c_{11} \equiv \gamma\sinh\beta\sum_{n=1}^{\infty}[\gamma\sinh(n\beta) + \sinh[(n-1)\beta]]^{-1},$$

$$c_{12} \equiv -\gamma\frac{\sinh\beta}{(1+\gamma)r}\sum_{n=1}^{\infty}[\sinh(n\beta)]^{-1},$$

$$c_{22} \equiv \gamma\sinh\beta\sum_{n=1}^{\infty}[\sinh(n\beta) + \gamma\sinh[(n-1)\beta]]^{-1}.$$

Dividing interaction energy (1) by $k|q_1q_2|/(R_1 + R_2)$, we obtain the dimensionless expression

$$W = \frac{1+\gamma}{2\alpha}\frac{\alpha^2 c_{11} - 2\alpha c_{12} + c_{22}}{c_{11}c_{22} - c_{12}^2}. \qquad (3)$$

The interaction force F is made dimensionless by dividing by the force $k|q_1q_2|/(R_1 + R_2)^2$ of Coulomb interaction between the point charges q_1 and q_2 separated by the center distance of the touching spheres:

$$F_r = -\frac{\partial W}{\partial r} = -\frac{r(1+\gamma)^3}{2\alpha\gamma\sinh\beta}$$
$$\times\left\{\frac{(\alpha^2 c'_{11} - 2\alpha c'_{12} + c'_{22})(c_{11}c_{22} - c_{12}^2)}{(c_{11}c_{22} - c_{12}^2)^2}\right. \qquad (4)$$
$$\left. - \frac{(\alpha^2 c_{11} - 2\alpha c_{12} + c_{22})(c'_{11}c_{22} + c_{11}c'_{22} - 2c_{12}c'_{12})}{(c_{11}c_{22} - c_{12}^2)^2}\right\}.$$

The (primed) derivatives of the capacitive coefficients with respect to β are defined by the formulas

$$c'_{11} \equiv -\sinh\beta\sum_{n=1}^{\infty}$$
$$\times\frac{\cosh\beta[n + \gamma(n-1)\cosh\beta] - \gamma(n-1)\sinh(n\beta)\sinh\beta}{[\sinh(n\beta)[1+\gamma\cosh\beta] - \gamma\cosh(n\beta)\sinh\beta]^2}$$
$$+ c_{11}\coth\beta,$$

$$c'_{22} \equiv -\sinh\beta\sum_{n=1}^{\infty}$$
$$\times\frac{\cosh\beta[\gamma n + (n-1)\cosh\beta] - (n-1)\sinh(n\beta)\sinh\beta}{[\sinh(n\beta)[\gamma + \cosh\beta] - \cosh(n\beta)\sinh\beta]^2}$$
$$+ c_{22}\cosh\beta,$$

$$c'_{12} \equiv c_{12}\coth\beta - c_{12}\frac{\sinh\beta}{(1+\gamma)^2 r^2}$$
$$+ \frac{\sinh\beta}{(1+\gamma)r}\sum_{n=1}^{\infty}\frac{n\cosh(n\beta)}{\sinh^2(n\beta)}.$$

For the computer analysis of the expressions for the energy, (3), and force, (4), we use the substitution proposed in [10]. Introducing the new variable $z \equiv \exp(-\beta)$ and setting $\cosh(b) \equiv y$, we obtain

$$c_{11} \equiv 2\gamma\sqrt{y^2-1}\sum_{n=1}^{\infty}$$
$$\times\frac{z^n}{(1-z^{2n})[(\gamma+y) - \sqrt{y^2-1}(1+z^{2n})/(1-z^{2n})]},$$

$$c_{22} \equiv 2\gamma\sqrt{y^2-1}\sum_{n=1}^{\infty} \qquad (5)$$
$$\times\frac{z^n}{(1-z^{2n})[(1-\gamma y) - \gamma\sqrt{y^2-1}(1+z^{2n})/(1-z^{2n})]},$$

$$c_{12} \equiv -\frac{2\gamma\sqrt{y^2-1}}{r(1+\gamma)}\sum_{n=1}^{\infty}\frac{z^{2n}}{1-z^{2n}}.$$

3. In Fig. 1, the surface of the dimensionless energy W of electrostatic interaction between the charged conducting spheres is shown in the space of the dimensionless parameter $r \equiv l/(R_1 + R_2)$ and the charge ratio $\alpha \equiv q_1/q_2$ for the ratio of their radii $\gamma = 0.3$. This result is calculated by (3) in view of (5). It is seen that the interaction energy of the like-charged spheres ($\alpha > 0$) may be both positive and negative, whereas for the oppositely charged spheres ($\alpha < 0$), the sign of the interaction energy remains unchanged. This fact seems to be evi-

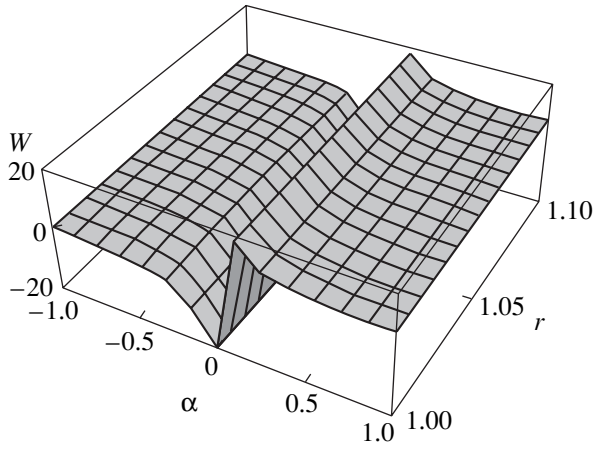


Fig. 1. Dimensionless electrostatic interaction energy W of charged conducting spheres as a function of the charge ratio α and dimensionless center distance r. $\gamma = 0.3$.

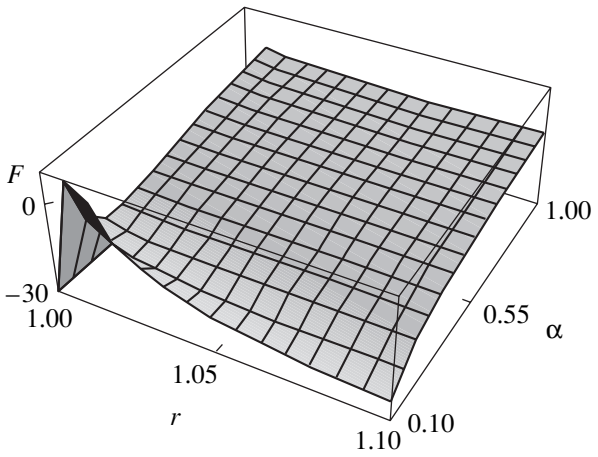


Fig. 2. Dimensionless force F acting on either like-charged sphere vs. dimensionless center distance r and the charge ratio α. $\gamma = 0.3$.

dent, since the energy of polarization interaction for charged conducting bodies is negative regardless of the signs of their charges. The same is true for the energy of Coulomb interaction between unlike charges. For like-charged bodies, the energy of Coulomb interaction is positive. The addition of this positive energy to the negative energy of polarization interaction, which rapidly increases with decreasing spacing between bodies, results in the negative total energy.

In Fig. 2, the surface of the dimensionless force acting on either sphere is shown in the space of the parameters r and α for $\gamma = 0.3$. The like-charged spheres are seen to be attracted to each other almost throughout the range of the parameters r and α. The exception is the narrow intervals of small values of the charge and spacing ratios where the repulsive forces act on the like-charged spheres. The interval of small spacings where the repulsive forces dominate corresponds to the case when the charge ratio is equal to the radius ratio squared (for details, see [10]). In this case, the fields of the self-charges on the surfaces equal each other and the like-charged spheres repel at arbitrarily small spacings.

The considerations concerning Fig. 2 are also illustrated by Figs. 3a–3c, where the dimensionless force F acting on either sphere is plotted as a function of the charge ratio for $\gamma = 0.3$ and various spacings r. From Fig. 3a, it follows that at a very small spacing ($r = 1.0001$) between the like-charged spheres, the acting force is an attractive force. Its value, which decreases with increasing α, is large for small α, namely, five or more times as great as the Coulomb force for the point charges q_1 and q_2 separated by a distance equal to the center distance of the touching spheres (this follows from the way used to make the force dimensionless). Over the distance $r = 1.01$ (Fig. 3b), as α increases, the repulsive force acting on the spheres first drops to zero, changes the sign, and becomes attractive. Then, it peaks

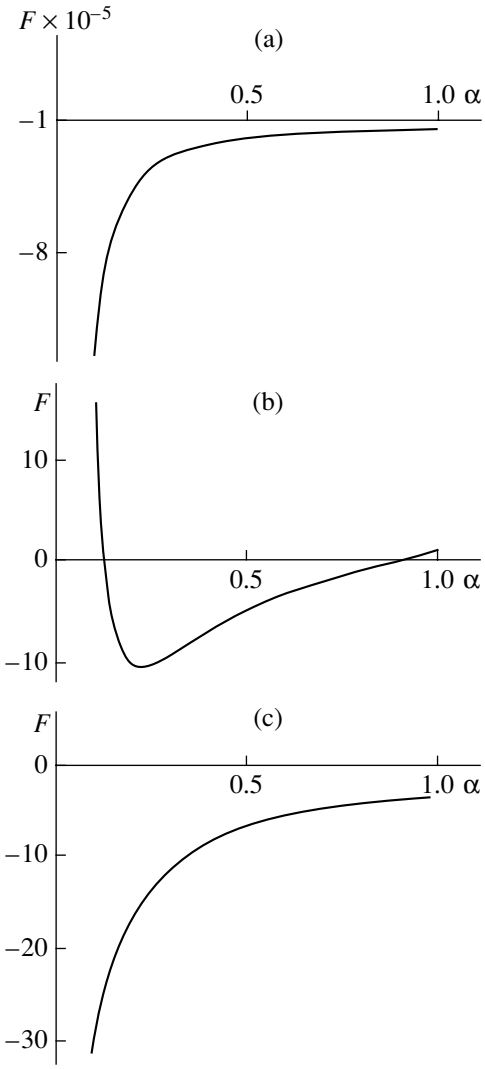


Fig. 3. Dimensionless force F acting on either like-charged sphere as a function of the charge ratio α at $\gamma = 0.3$. r = (a) 1.0001, (b) 1.01, and (c) 1.1.

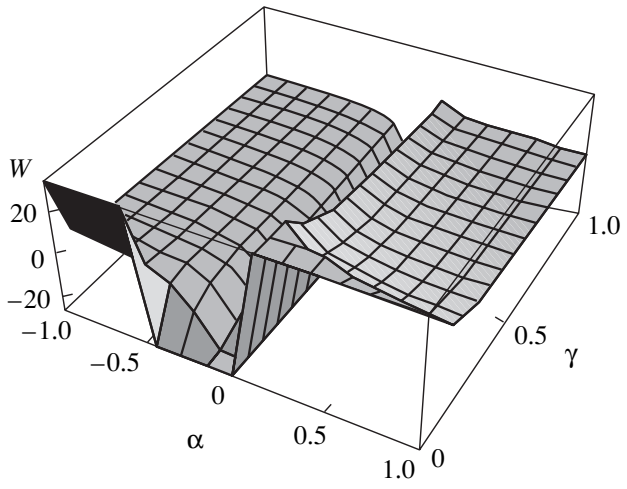


Fig. 4. Dimensionless energy *W* vs. ratios of the charges, α, and radii, γ. *r* = 1.01.

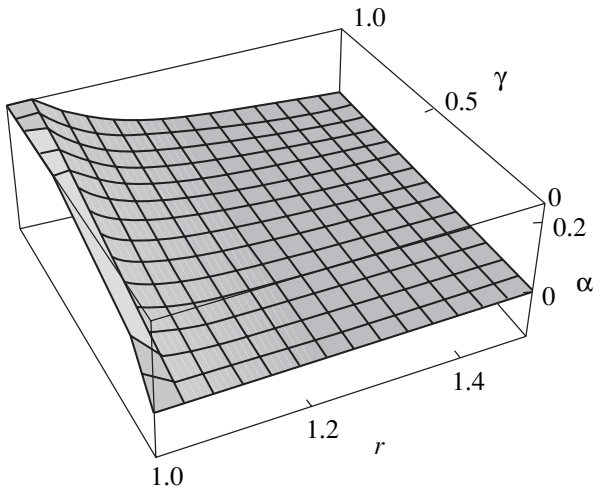


Fig. 5. Surface separating the ranges of the parameters α, γ, and *r* where the repulsive and attractive forces dominate.

and drops to zero again as α increases further. Finally, it changes the sign once more and becomes again repulsive. For $r = 1.1$ (Fig. 3c), in the whole range of the parameters, the like-charged spheres experience only the attractive force.

The large attractive forces arising between the like-charged particles at small spacings (Fig. 3a) correspond to high electrostatic field strengths. Under real conditions, this will result in air (or vacuum) breakdown in the gap between the like-charged particles approaching each other, causing the charge redistribution among them. The direction of charge transfer during the discharge between the like-charged particles depends on how the electrostatic energy of the system will decrease. In Fig. 4, the dependence of the dimensionless energy *W* of electrostatic interaction between the two conducting spheres on the ratios of their charges, α, and radii, γ, is shown. In a closed system, the only processes that can spontaneously take place are those resulting in a decrease in the potential energy (in our case, the electrostatic interaction energy). From Fig. 4, it is easily seen that for $\alpha > 0$ and a constant ratio of the sphere radii ($\gamma = \text{const}$), the potential energy would decrease only with increasing charge ratio α.

In the case of approaching charged drops (instead of solid spheres), the surface electrostatic instability of one of them occurs at field strengths even lower than those required for air breakdown. This process is accompanied by the ejection of a liquid jet towards the other drop, as has been observed in experiments [11–15], which may cause the coagulation of like-charged drops. Figure 4 can be a help in analyzing the direction of liquid charge and mass transfer when the electrohydrodynamic instability of like-charged drops takes place in unipolar charged liquid-drop systems. Since the point $\alpha = 1$, $\gamma = 1$ in Fig. 4 is the minimum of the potential energy, the potential energy of a closed system consisting of two spheres will decrease with increasing ratios of the charges, α, and radii, γ; i.e., the tendency $\alpha \longrightarrow 1$ and $\gamma \longrightarrow 1$ will be observed.

Figure 5 shows the surface (on which the force defined by relation (4) vanishes) separating the ranges of the parameters where the like-charged spheres repel (below the surface) and attract (above the surface) each other in the space of the parameters α, γ, and *r*.

CONCLUSION

When like-charged conducting particles (drops) approach, the electrostatic repulsion forces are replaced by the forces of polarization attraction at small spacings between the particles (drops). Under these conditions, the electrostatic field between the drops may cause air (or vacuum) breakdown. As a result, charge transfer between solid particles will take place. In the case of the drops, this process will be accompanied by liquid transfer due to the instability of the charged surface near the tops of drops.

REFERENCES

1. M. H. Davis, Q. J. Mech. Appl. Math. **176**, 499 (1964).
2. A. I. Grigor'ev and S. O. Shiryaeva, J. Phys. D **23**, 1361 (1990).
3. A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61** (3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].
4. E. I. Mukhina and A. I. Grigor'ev, Zh. Tekh. Fiz. **62** (2), 18 (1992) [Sov. Phys. Tech. Phys. **37**, 122 (1992)].
5. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
6. V. A. Koromyslov, A. I. Grigor'ev, and S. O. Shiryaeva, Zh. Tekh. Fiz. **68** (8), 31 (1998) [Tech. Phys. **43**, 904 (1998)].

7. A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (5), 22 (2000) [Tech. Phys. **45**, 543 (2000)].
8. I. E. Mazets, Zh. Tekh. Fiz. **70** (10), 8 (2000) [Tech. Phys. **45**, 1238 (2000)].
9. W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill, New York, 1950; Inostrannaya Literatura, Moscow, 1954).
10. V. A. Saranin, Usp. Fiz. Nauk **169**, 453 (1999).
11. W. A. Macky, Proc. R. Soc. London, Ser. A **133**, 565 (1931).
12. J. Latham and I. W. Roxburgh, Proc. R. Soc. London, Ser. A **295**, 84 (1966).
13. J. Latham, in *Planetary Electrodynamics: Proceedings of the 4th International Conference on the Universal Aspects of Atmospheric Electricity, 1968*, Ed. by S. C. Coroniti and J. Hughes (Gordon and Beach, New York, 1969), Vol. 1, p. 345.
14. A. H. Miler, C. E. Shelden, and W. R. Atkinson, Phys. Fluids **8**, 1921 (1965).
15. J. D. Sartor and C. E. Abbott, J. Geophys. Res. **73**, 6415 (1968).

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 1, 2002, pp. 132–134. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 136–138.

BRIEF COMMUNICATIONS

On the Reality of Supraluminal Group Velocity and Negative Delay Time for a Wave Packet in a Dispersion Medium

N. S. Bukhman

Samara State Academy of Architecture and Civil Engineering, Samara, Russia

Received May 8, 2001

Abstract—It is shown that, in the case of the supraluminal group velocity of a wave packet in a dispersion medium, a wave packet with a smooth (analytical) envelope does propagate with a supraluminal velocity. In the case of a negative group velocity, the signal maximum does arrive at the detector earlier than at the transmitter. These facts are consistent with both the finiteness of the velocity of light in free space for information transfer (in the case of supraluminal propagation velocity) and the principle of causality (in the case of negative delay time). Basically, the effect of negative delay time may be employed for predicting an observable effect. © *2002 MAIK "Nauka/Interperiodica".*

It is well known that, in a time-dispersion medium, a wave packet with a smoothly varying envelope (i.e., with a narrow frequency spectrum) propagates with the so-called group velocity. Generally, the group velocity may be both greater and smaller than the velocity of light in free space (or even be negative) [1, 2]. The supraluminal or negative group velocity of a signal is sometimes considered as a formal result (see, e.g., [2]). Nevertheless, the motion of the wave packet maximum with a supraluminal velocity in a nonlinearly amplifying medium has been studied both theoretically and experimentally in a number of papers [3, 4]. In particular, this effect in a nonlinear medium with gain saturation has been related to the preferred amplification of the leading edge of a signal compared with the trailing one.

In this paper, we point to the following aspects.

(i) This effect (signal propagation with a supraluminal velocity) takes place for weak signals and, in the linear approximation, occurs as an inevitable result of the supraluminal group velocity of a wave at a certain carrier frequency. This effect is not an inherent feature of electromagnetic waves in space: it may occur when a signal of any nature propagates through an arbitrary linear filter [5] provided that the signal spectrum is narrow and the carrier frequency is properly selected.

(ii) The effect takes place in both a linear amplifying medium (at the wing of an amplification spectrum line) and a linear absorbing medium (at the center of an absorption spectrum line).

(iii) In the case of sufficiently strong dispersion, the signal delay time may become not only less than the time it takes for light to propagate from the transmitter to the detector in free space, but even negative. This means that the signal is detected before it is transmitted. The aforesaid allows one to draw the conclusion on the fundamental possibility of "advanced reflection" in the inorganic nature.

In order to illustrate the above statements, consider a high-frequency signal $E(z, t)$ with the carrier frequency ω_1 and the complex envelope $A(z, t)$ [1, 2] that propagates along the z axis. We assume that the signal frequency ω_1 is close to the frequency ω_0 of a spectral line of the medium. If the signal is narrow-band (the spectral width of the signal is small compared to the carrier frequency ω_1), we have the following evident relations:

$$E(z, t) = A(z, t)\exp(-i\omega_1 t) + A^*(z, t)\exp(i\omega_1 t),$$

$$A(z, t) = \int_{-\infty}^{+\infty} A(z, \Delta\Omega)\exp(-i\Delta\Omega t)d\Delta\Omega, \qquad (1)$$

$$A(z, \Delta\Omega) = E(z, \omega), \quad \omega = \omega_1 + \Delta\Omega, \quad \omega_1 = \omega_0 + \Omega_0.$$

Here, $E(z, t)$ is the high-frequency signal; $A(z, t)$ and $A(z, \Delta\Omega)$ are the low-frequency complex envelope of the signal and its spectrum, respectively; ω_0 is the center frequency of the spectral line; and Ω_0 is the shift of the carrier frequency with respect to the center of the spectral line. We assume that the signal propagates in a gas with a refractive index $m(\omega) = 1 + 2\pi\chi(\omega)$, where $\chi(\omega)$ is the complex dielectric susceptibility of the gas. Then, the complex transfer function for a layer of width z is $F(z, \omega) = \exp(ikn(\omega)z)$, where $k = \omega/c$. Let us introduce the amplitude gain for light at the center of the spectral line, $\alpha_0 \equiv 2\pi i k_0 \chi(\omega_0)$ ($k_0 \equiv \omega_0/c$), and the complex spectral form factor $g(\Omega) \equiv 2\pi i k \alpha_0^{-1} \chi(\omega_0 + \Omega)$ normalized to unity at the center of the spectral line ω_0. Then, the transfer function of the layer can be represented in the form $F(z, \omega) = \exp(ikz)\exp(\xi g(\Omega))$, where $\xi \equiv \alpha_0 z$ is the optical thickness of the layer. In the case

1063-7842/02/4701-0132$22.00

of an absorption spectral line, the gain α_0 is negative (as well as the optical thickness ξ).

Assume that the complex envelope of the signal at the initial point $z = 0$ is defined by the function $A^{(0)}(t) \equiv A(0, t)$. Then, in the cross section z, we have

$$A(z, t) = \int_{-\infty}^{+\infty} A^{(0)}(\Delta\Omega)F(z, \omega)\exp(-i\Delta\Omega t)d\Delta\Omega. \quad (2)$$

Let the signal spectrum be concentrated in the vicinity of the carrier frequency ω_1. Leaving the linear terms in the Taylor expansion of the exponent in (2) (i.e., considering the first order of the classical dispersion theory [1, 2]), one can readily obtain

$$A(z, t) = \exp(ik_1 z + \xi g(\Omega_0))A^{(0)}(t - \tau(z)), \quad (3)$$

from (2). In (3), the complex delay time τ is defined as

$$\tau(z) \equiv \tau_{vac} + \tau_r + i\tau_i, \quad \tau_{vac} \equiv z/c,$$
$$\tau_r \equiv \xi\frac{\partial(\mathrm{Im}g(\Omega_0))}{\partial\Omega_0}, \quad \tau_i \equiv -\xi\frac{\partial(\mathrm{Re}g(\Omega_0))}{\partial\Omega_0}. \quad (4)$$

The result obtained differs from the conventional first-order expression of dispersion theory [1, 2] only by taking into account the imaginary part of the delay time of a wave packet. Note that in this case (in contrast to the nonlinear case [3, 4]), the delay time of the signal is independent of its duration, power, and envelope shape, but significantly depends on the shift of the carrier frequency with respect to the center of the spectral line.

For definiteness, we will consider a Gaussian wave packet with a duration T ($A^{(0)}(t) = \exp(-t^2/T^2)$) and the Lorentzian profile of a spectral line [6] having a width $\Delta\Omega_{1/2}$ and a coherence time $\tau_l \equiv 2/\Delta\Omega_{1/2}$, $g(\Omega) = (1 - i2\Omega/\Delta\Omega_{1/2})^{-1}$. In this case, the time dependence of the field intensity $I(z, t) \equiv |A(z, t)|^2$ for various values of the longitudinal coordinate z is given by [instead of (3)]

$$I(z, t) = I_0(z)I_G(z)\exp(-2(t - \Delta t)^2/T^2),$$
$$I_0(z) \equiv \exp(2\xi/(1 + x_0^2)),$$
$$I_G(z) \equiv \exp[8(\tau_l/T)^2\xi^2 x_0^2(1 + x_0^2)^{-4}], \quad (5)$$
$$\Delta t \equiv z/c + \xi\tau_l(1 - x_0^2)(1 + x_0^2)^{-2}, \quad x_0 \equiv \tau_l\Omega_0,$$

where $I_0(z)$ is the conventional factor of exponential increase (decrease) in the intensity of a monochromatic wave with a frequency $\omega_1 = \omega_0 + \Omega_0$ in an amplifying (absorbing) medium, $I_G(z)$ is the additional (with respect to the monochromatic wave) factor increasing the intensity of the Gaussian packet, Δt is the real packet delay time, and x_0 is the normalized detuning of the carrier frequency from the center of the line.

In the case of the Gaussian shape of the signal, the imaginary part of the delay time results only in an additional (with respect to the monochromatic wave) increase in the signal amplitude.

If $x_0^2 > 1$ and $\xi > 0$ (i.e., at the wings of the amplification spectral line) or $x_0^2 < 1$ and $\xi < 0$ (i.e., near the center of the absorption spectral line), the wave packet propagates with a supraluminal velocity (the delay time $\Delta t < z/c$). This result is in no way inconsistent with the finiteness of the velocity of light in free space for information transfer. Indeed, the discontinuities of the signal envelope always propagate with the velocity of light [1]. Therefore, the supraluminal velocity of the time envelope of a signal cannot be used for information transfer with a velocity exceeding the velocity of light (for details, see [3, 4]).

Relations (5) show that, at a sufficiently large amplification (absorption) coefficient of the medium, the signal delay time is negative ($\Delta t < 0$) for any length of the path z and increases (in magnitude) with increasing the path length. This means that the time dependence of the signal at the detector is advanced with respect to that at the transmitter. Note that this advance may exceed the characteristic duration of the signal. The effect of advanced signal arrival at the detector with respect to the transmitter can, in principle, be used to predict the future. Nevertheless, this result is consistent with the principle of causality for the same reasons that explain why a supraluminal signal velocity is not in conflict with the finiteness of the velocity of light. The advanced signal arrival at the detector should be considered as the natural prediction of the time dependence of a signal in the future on the basis of the time dependence of the signal in the past available at the transmitter.

In order to explain this, it is useful to consider the transmission of a signal with the cut trailing edge. In this case, it is the finiteness of the velocity of light in free space (vacuum) for information transmission that results in the reconstruction of the trailing edge, which has not been completely transmitted. Indeed, in the case of supraluminal group velocity of a signal, the information about its abrupt turn-off propagates more slowly (with the velocity of light) than the signal itself. Therefore, the trailing edge of the signal is detected even if it is not transmitted at all (the transmission is abruptly terminated).

The analytical results are illustrated by numerically calculating the propagation of a signal with the duration $T = 5\tau_l$ and the normalized carrier frequency shift $x_0 = 2$ at $\xi = 0$, 10, and 20 (amplifying medium, wing of amplification line) (Fig. 1). The real signal amplitude $A_n(z, t) \equiv \sqrt{I(z, t)/I_0(z)}$ normalized to the amplitude of a plane wave with the same frequency is plotted on the ordinate; time in the co-moving coordinate system (moving with the velocity of light in free space), on the abscissa. The calculations were made for the uncut signal (Fig. 1a), as well as for the signal for which transmission was abruptly terminated at $t = 0$ (Fig. 1b). The solid lines correspond to the results of numerical calculations, the dashed lines show the results obtained by

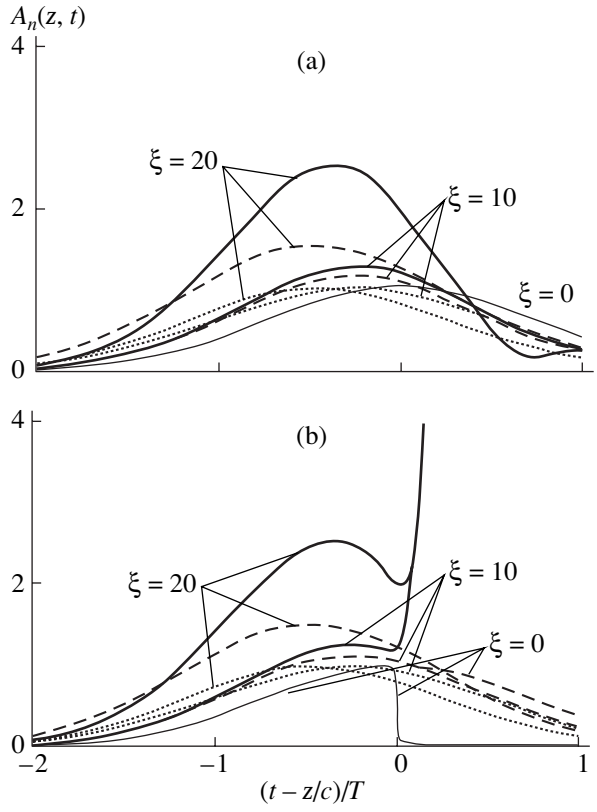


Fig. 1.

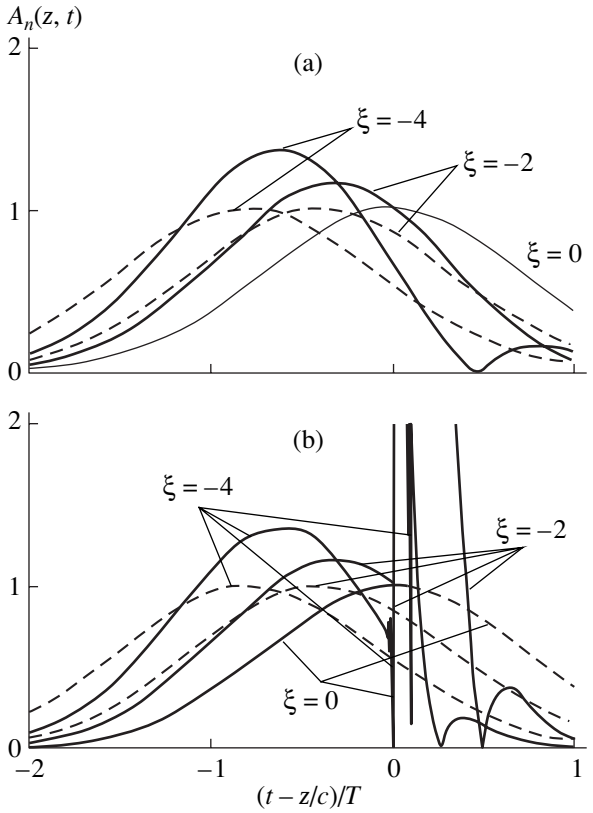


Fig. 2.

formula (5), and the dotted lines show the results calculated by formula (5) disregarding the factor $I_G(z)$.

Similar calculations for an absorbing medium are shown in Figs. 2a and 2b. They were performed for a signal with the duration $T = 5\tau_l$ and the normalized shift of the carrier frequency $x_0 = 0$ at $\xi = 0$, –2, and –4 (absorbing medium, the center of the absorbing line). As is seen, the signal propagates with the supraluminal velocity. The poor accuracy of the analytical results found in the first-order of the classical dispersion theory is related to the wide signal spectrum.

Comparing Figs. 1 and 2, one can easily see that a part of the trailing edge is successfully detected far away from the transmitter (even if it has not been transmitted at all). The untransmitted part of the signal is detected until the information that the signal is not transmitted arrives at the detector (with the velocity of light). Then, the sharp increase in the signal amplitude is observed, since the spectrum of the partially cut signal substantially widens (compared to the "expected" one). A substantial part of the signal is concentrated either at the center of the amplification spectral line (in the case of an amplifying medium) or outside the absorption spectral line (in the case of an absorbing medium). In any case, the signal detected may be considered as the prediction of the transmitted one, the prediction being made by the medium under the assumption about the analyticity of the signal envelope. This feature of the signal detected allows one to admit the fundamental possibility (under certain conditions) of advanced reception of information for an observable process.

ACKNOWLEDGMENTS

The author is grateful to A.A. Rukhadze and S.V. Bulanov for the valuable discussions.

REFERENCES

1. A. I. Akhiezer and I. A. Akhiezer, *Electromagnetism and Electromagnetic Waves* (Vysshaya Shkola, Moscow, 1985).
2. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *The Theory of Waves* (Nauka, Moscow, 1979).
3. N. G. Basov, R. V. Ambartsumyan, *et al.*, Dokl. Akad. Nauk SSSR **165**, 58 (1965) [Sov. Phys. Dokl. **10**, 1039 (1966)].
4. P. G. Kryukov and V. S. Letokhov, Usp. Fiz. Nauk **99**, 169 (1969) [Sov. Phys. Usp. **12**, 641 (1970)].
5. D. V. Vasil'ev, M. R. Vitol', Yu. N. Gorshenkov, *et al.*, *Radio Engineering Circuits and Signals* (Radio i Svyaz', Moscow, 1982).
6. D. N. Klyshko, *Physical Foundations of Quantum Electronics* (Nauka, Moscow, 1986).

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 1, 2002, pp. 135–138. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 139–142.

BRIEF COMMUNICATIONS

Low-Pressure Electrode Microwave Discharge in Argon and in an Ar–O_2 Mixture

Yu. A. Lebedev and M. V. Mokeev

Topchiev Institute of Petrochemical Synthesis, Russian Academy of Sciences, Leninskiĭ pr. 29, Moscow, 117912 Russia

e-mail: lebedev@ips.ac.ru

Received May 30, 2001

Abstract—The phenomenological description of an electrode microwave discharge in argon and in an Ar–O_2 mixture under a pressure between 1 and 30 torr and an incident power of below 300 W is presented. The results of probe measurements are reported. It is shown that the discharge can be initiated at the tip of an exiting rod antenna and takes different forms, for instance, spherical or the form of moving glowing filaments. In our case, the spatial structure of a spherical discharge is more uniform than that of discharges excited earlier in molecular gases. The addition of oxygen raises the power absorbed by the plasma (if the incident power is kept constant) and brings the discharge structure closer to that in molecular gases. © *2002 MAIK "Nauka/Interperiodica".*

The study of electrode microwave discharges is a new branch of investigation in the field of gas discharge. Information on this question is nearly absent in the literature. This is due to the tendency for a removal of the electrodes as a contamination source, from the discharge chamber. A theory of these discharges is lacking, and experimental data on their properties and parameters are being collected only now. Electrode discharges in molecular gases have been experimentally investigated in a number of papers [1–5]. These experiments have allowed researchers to elaborate the general ideas of physical processes occurring in electrode microwave discharges. These processes are responsible for the structure observed and are basic to the experimental relationships obtained. Estimates show that the electric field strength outside the glow region is insufficient for discharge maintenance, so that the plasma in this region decays. The glow region is, in essence, the region of a self-sustained discharge, where the energy is absorbed. It has a nonunifom spherical structure with a bright near-electrode layer. As has been shown, such discharges can be employed in plasma chemistry [6, 7]. In this paper, we investigated the discharges initiated in argon and in a 84 vol % Ar + 16 vol % O_2 mixture.

The discharge system was similar to that described in [1–5]. The discharge was initiated in the evacuated metal chamber. The electromagnetic energy (2.45 GHz, incident power $P_{in} \le 300$ W) was delivered to the chamber through a stainless steel rod of diameter 6 mm. The incident power and the power reflected by the discharge were measured with a directional coupler. The matching of the discharge chamber to the generator could be varied by using shorting plungers inserted in the feeding waveguide. The discharge was initiated at the end of the rod. The argon pressure in the chamber was 1–30 torr, and the argon flow rate was no more than 100 sccm.

The discharge was visualized with a Panasonic NV-M3000EN video camera (the exposure times between 1/20 and 1/8000 s) with the subsequent digital processing of the images (see Fig. 1). The plasma parameters shown in Figs. 2a and 2b were determined with the method of double electric probes [1, 2]. The ion saturation current I_i (see Fig. 2a) was measured at a probe voltage of 50 V and averaged over its values calculated on both branches of the double-probe *I–V* characteristic. The error in determining the ion current was about 10%. The charged particle density N_i can be estimated from the saturation ion current with the well-known relationship [8]

$$I_i = 0.61 N_i e S_p v_i, \quad (1)$$

where e is the electron charge, $S_p = \pi dl = 4.7 \times 10^{-7}$ m^2, $v_i \approx 10^4 (T_e/A)^{1/2}$ is the mean velocity of ions with allowance for their acceleration in the probe presheath, and A is the mass number of the ion.

Let Ar^+ ions dominate in the argon plasma. Then, to estimate the charged particle density ($n_e \approx n_i$), we can take advantage of the relationship $N_i[\text{cm}^{-3}] \approx 7 \times 10^9 I_i[\mu\text{A}]$ (we assume here that $T_e \approx 3$ eV). The electron temperature was conventionally determined from the double-probe *I–V* characteristic near the floating potential (see, for instance, [8]). The data shown in Fig. 2 characterize the time-averaged properties of the discharge, since it often represents a set of moving filamentary structures (Fig. 1) and the time it takes for the probe *I–V* curve to be measured exceeds that within which the discharge structure changes. The basic

1063-7842/02/4701-0135$22.00

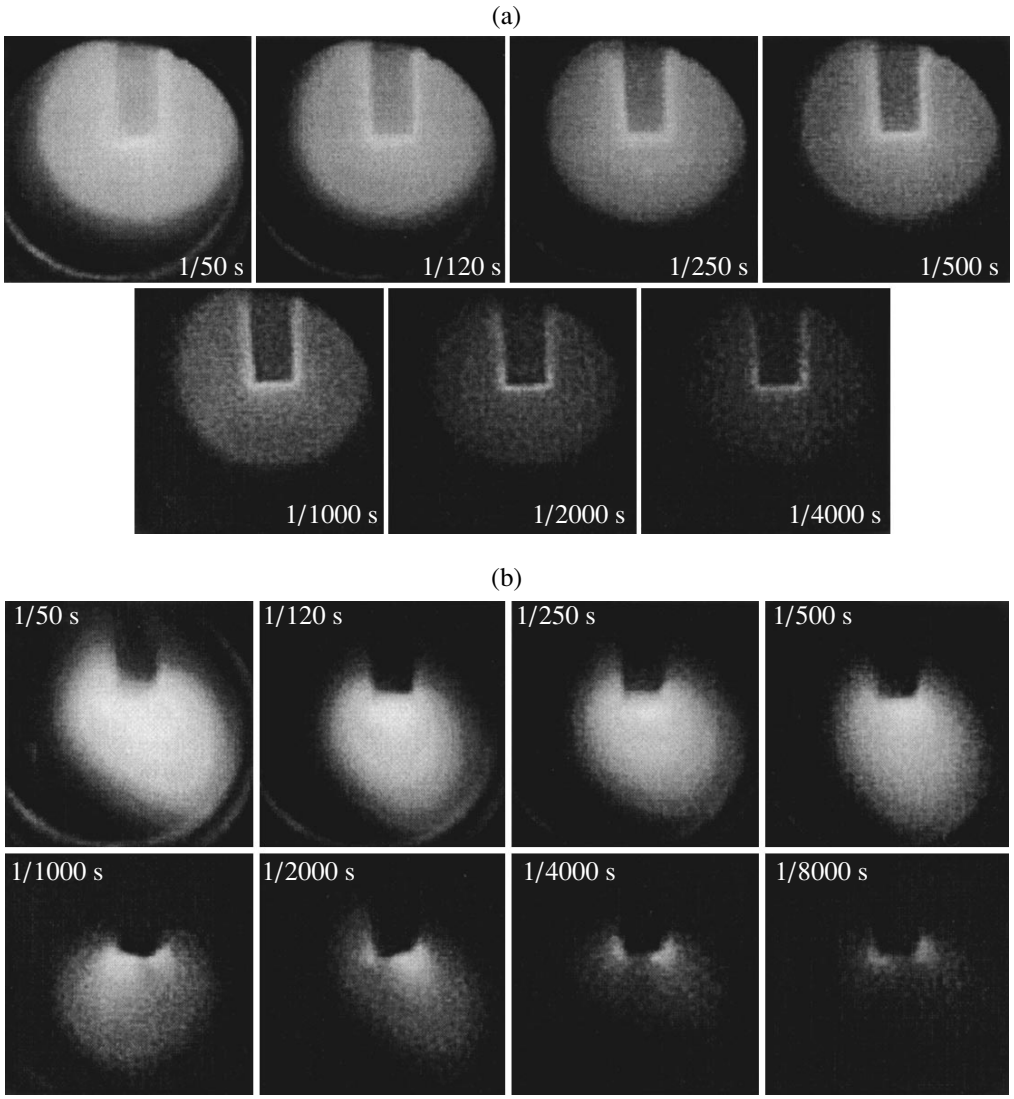


Fig. 1. Electrode discharge in argon at various exposure times under the pressure (a) 1, (b) 10, (c) 20, and (d) 30 torr (the dimension of the plasma region can be found by comparing with the image of the cylindrical electrode 6 mm in diameter).

results obtained in our work can be summarized as follows.

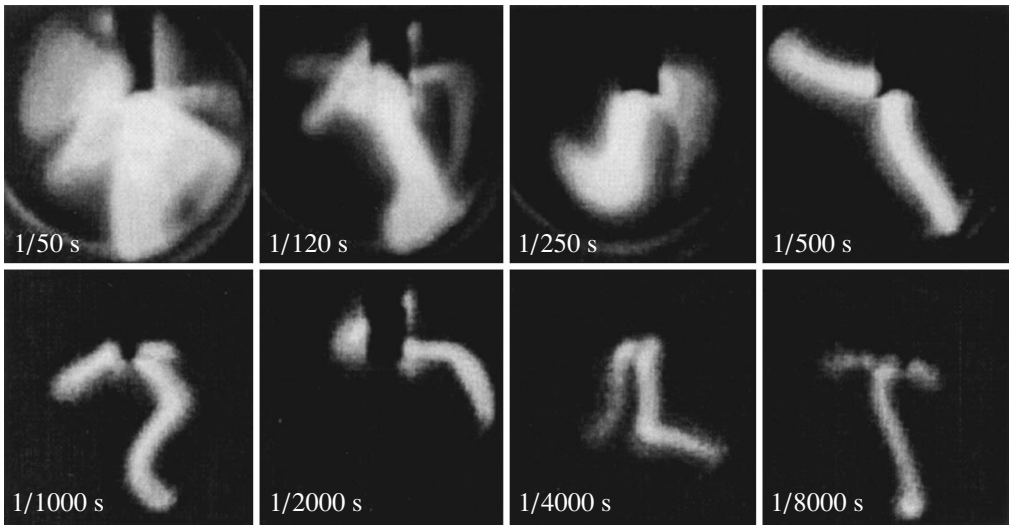
There are two forms of the discharge. When the pressure is no more than 10 torr, it takes the spherical form (Figs. 1a, 1b). With an increase in the pressure, the discharge is comprised of moving filaments originating at the electrode (Figs. 1c, 1d). The filamentary structure becomes developed at higher pressures. When observing the discharge with various exposures times at high pressures, one can distinguish the moving filaments in the seemingly diffuse structure.

In the case of the spherical discharge, the glow intensity is high at the antenna, monotonically falling towards the discharge periphery.

If the microwave power is high, the discharge is initiated at the point where the antenna enters the chamber (the taper of the coaxial line); then, the discharge can be shifted toward the end of the antenna by varying the position of the plungers in the waveguide. As a rule, the discharge is excited only at a certain combination of the position of the plungers and the power level. A change in either factor leads to a change in the other. The system shows the resonant properties.

The spherical discharge in argon at the end of the antenna is excited when the level of reflection significantly exceeds that in molecular gases. It is due to the fact that the reactive component of the plasma impedance in inert gases is much higher than in molecular gases. The reflection decreases only after an additional discharge has been initiated at the point where the antenna enters the chamber; this discharge absorbs the power almost completely (its glow intensity grows with power).

(c)



(d)

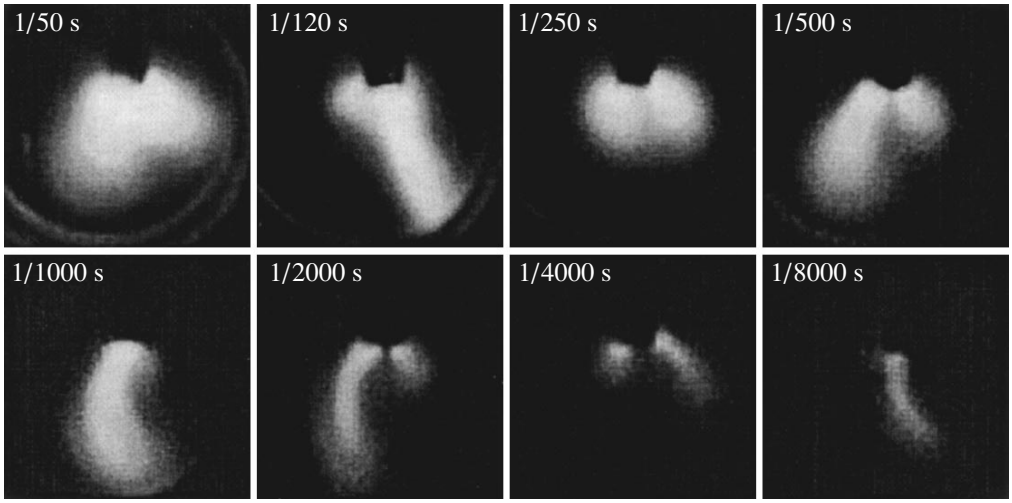


Fig. 1. (Contd.)

The addition of a molecular gas (oxygen) stabilizes the discharge and makes its boundary more distinct. If the incident power is kept constant, the electron temperature (Fig. 2b, curve *3*) and the charged particle density (Fig. 2a, curve *4*) grow because the power absorbed in the plasma grows. The discharge structure approaches that in molecular gases [4, 5].

The electron density at the discharge axis is above the critical value (at a frequency of 2.45 GHz, $n_{ec} \approx 7 \times 10^{10}$ cm^{-3}).

It has been shown [2] that the probe measurements of the electron temperature around a spherical electrode discharge in hydrogen can be explained by the presence of a surface electromagnetic wave with the field strength exponentially dropping along the normal to the discharge boundary. This is observed at pressures when the radial variation of the electron temperature is governed by the heating of the electrons in the electromagnetic field and the typical scale of collisional electron-energy dissipation is much smaller than the characteristic scale of field variation outside the plasma region. In hydrogen, this took place at pressures of 1–15 torr. In argon, where the relative electron energy loss, $\delta \sim 10^{-4}$, is much less than in molecular gases, $\sim 10^{-3}$, the effect should show up at higher pressures. An example is curve *3* in Fig. 2b: the exponential decline in the electron temperature starts only at 8 torr. In this case, the electric field strength can be estimated with the Boltzmann kinetic equation, as in [2]. It makes it possible to calculate the field strength corresponding to the electron temperatures measured and can be solved numerically [9]. For an argon plasma, the estimation is more difficult, because the "core" of the electron energy distribution strongly depends on the electron density. This

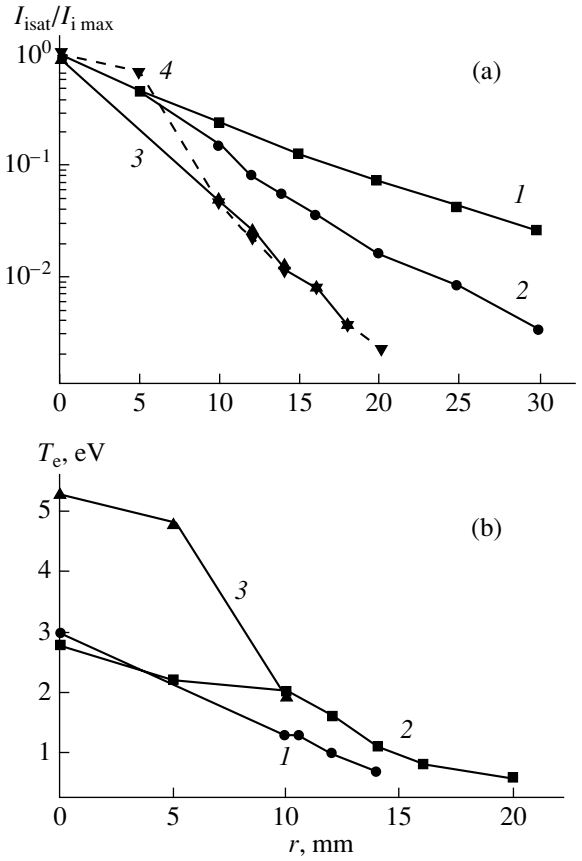


Fig. 2. (a) Radial distribution for the ion saturation current of the probes normalized to its peak value $I_{i\,max}$ for various pressures, incident powers, and currents $I_{i\,max}$ in Ar: (*1*) 1 torr, 140 W, 24 μA; (*2*) 3 torr, 200 W, 80 μA; (*3*) 8 torr, 200 W, 92 μA; (*4*) 3 torr, 200 W, 140 μA for the 84% Ar + 16% O_2 plasma. (b) Radial distribution of the plasma electron temperature at an incident power of 200 W under the pressure (*1*) 8 and (*2*) 3 torr in the Ar plasma and (*3*) 3 torr in the 84% Ar + 16% O_2 plasma.

is because of the appreciable effect of electron–electron collisions. In the vicinity of the argon plasma, the electric field strength was estimated at 2.5–4 V/cm, which is significantly below the values obtained for hydrogen [2].

ACKNOWLEDGMENTS

This work was partially supported by the NWO (Netherlands) (grant no. 047.011.000.01).

We are grateful to A.A. Skovoroda for the technical assistance.

REFERENCES

1. L. Bardos and Yu. A. Lebedev, Fiz. Plazmy **24**, 956 (1998) [Plasma Phys. Rep. **24**, 891 (1998)].
2. L. Bardos and Yu. A. Lebedev, Zh. Tekh. Fiz. **68** (12), 29 (1998) [Tech. Phys. **43**, 1428 (1998)].
3. Yu. A. Lebedev, M. V. Mokeev, and A. V. Tatarinov, Fiz. Plazmy **26**, 293 (2000) [Plasma Phys. Rep. **26**, 272 (2000)].
4. Yu. A. Lebedev and M. V. Mokeev, Teplofiz. Vys. Temp. **38**, 381 (2000) [High. Temp. **38**, 338 (2000)].
5. L. Bardos and Yu. A. Lebedev, Teplofiz. Vys. Temp. **38**, 552 (2000) [High. Temp. **38**, 528 (2000)].
6. L. Bardos, H. Barankova, Yu. A. Lebedev, *et al.*, Diamond Relat. Mater. **6**, 224 (1997).
7. L. Bardos, H. Barankova, and Yu. A. Lebedev, in *Proceedings of the 42nd Annual Technical Conference of Society of Vacuum Coaters (SVC TC), Chicago, 1999*, Paper E-7.
8. *Plasma Diagnostics*, Ed. by W. Lochte-Holtgreven (Elsevier, New York, 1968; Mir, Moscow, 1971).
9. D. I. Aleksandrov, E. M. Dianov, L. M. Blinov, *et al.*, Fiz. Plazmy **12**, 1008 (1986) [Sov. J. Plasma Phys. **12**, 574 (1986)].

Translated by B. Malyukov

Technical Physics, Vol. 47, No. 1, 2002, p. 139. Translated from Zhurnal Tekhnichesko ĭ Fiziki, Vol. 72, No. 1, 2002, p. 143.

INFORMATION

Information on the Ostrovskiĭ Prize Competition

In 2001, the Ioffe Physicotechnical Institute (Russian Academy of Sciences) and the joint-stock bank Russia conducted the successive competition in honor of Yu.I. Ostrovskiĭ, the founder of the Russian school of holographic interferometry.

On April 12, 2001, the Awarding Commission of the Ostrovskiĭ Prize consisting of Yu.N. Denisyuk, Academician, Russian Academy of Sciences (Chairman); G.V. Ostrovskaya, Dr. Sci. (Phys.–Math.), Ioffe Physicotechnical Institute; D.I. Stasel'ko, Dr. Sci. (Phys.–Math.), All-Russia Research Center, Vavilov State Optical Institute; Yu.A. Tolmachev, Dr. Sci. (Phys.–Math.), St. Petersburg State University; and G.V. Dreĭden, Cand. Sci. (Phys.–Math.), Ioffe Physicotechnical Institute (Secretary) considered eight works submitted to the 2001 competition.

The first prize (two-thirds of the bonus fund) was unanimously awarded to I.A. Sokolov (Ioffe Physicotechnical Institute) for the work "Adaptive Photodetectors based on Dynamic Holographic Gratings in Photorefractive Crystals."

The second prize (one-third of the bonus fund) was awarded to E.A. Sokolova (All-Russia Research Center, Vavilov State Optical Institute) for the work "New-Generation Diffraction Gratings: The Theory, Production, and Application in Spectral Devices."

The Commission notes the high scientific level of the works submitted and thanks all authors participating in the competition.

The 2002 competition among the best works concerned with optical holography and holographic interferometry was announced. The bonus fund size is 20 000 rubles.

Works to the 2002 Ostrovskiĭ Prize competition should be directed (in three copies) to: G.V. Dreĭden (Secretary of the Commission), Ioffe Physicotechnical Institute, Russian Academy of Sciences, ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia. Deadline for submission: March 30, 2002

Phone (812)247-91-85
e-mail: Galina.Dreiden@pop.ioffe.rssi.ru

1063-7842/02/4701-0139$22.00

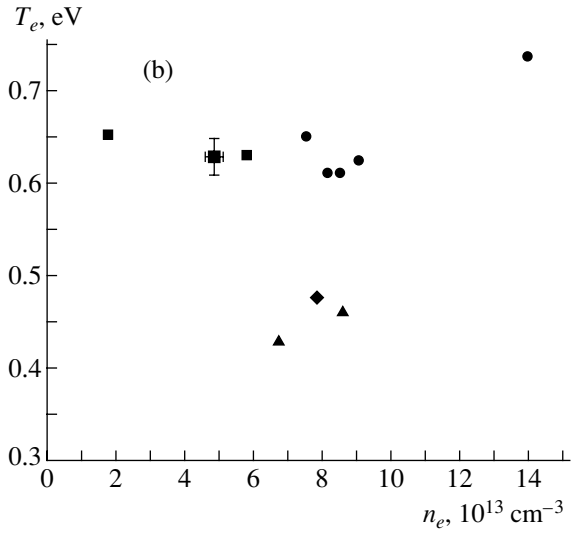
Technical Physics, Vol. 47, No. 1, 2002, p. 140.

ERRATA

Erratum: "Determination of the Negative Hydrogen Ion Concentration in a Cesium–Hydrogen Discharge" [*Technical Physics* 46 (8), 946 (2001)]

F. G. Baksht, V. G. Ivanov, S. I. Kon'kov, and S. M. Shkol'nik

Wrong values were indicated on the *y*-axis of Fig. 4b (page 949.) The correct version of this figure is given below.



1063-7842/02/4701-0140$22.00

Technical Physics, Vol. 47, No. 1, 2002, pp. 18–25. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 20–27.

Differential Characteristic of the Flow behind a Shock Wave

A. V. Omel'chenko

St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 198904 Russia

Received May 28, 2001

Abstract—Relationships between the derivatives on both sides of a discontinuity in a nonstationary shock wave moving with acceleration in a one-dimensional vortex flow of perfect gas are deduced. The problem of interaction between the shock wave and a weak discontinuity is solved based on these relationships. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The need for deriving relationships between the properties of a strong discontinuity, such as the acceleration or the curvature of a shock wave, and the derivatives of gasdynamic variables on both sides of the strong discontinuity has been associated largely with two problems. The first is the study of flows behind curved shock waves, and the second is the analysis of strong–weak discontinuity interaction. Early results in this field, which date back to late in the 1940s [1, 2], were concerned with the special case of a planar or an axisymmetric stationary curved shock wave. Later, they were extended for higher dimensionality problems [3, 4]. However, most relationships that related the derivatives on both sides of a strong discontinuity were awkward. Therefore, the problem of strong–weak discontinuity interaction in gas dynamics either was solved by the perturbation method or was considered as a special case of strong–strong discontinuity interaction [5].

In this work, we derive simple relationships between the derivatives on both sides of a nonstationary one-dimensional shock wave. Based on them, we attack the problem of shock wave interaction with counter and weak cocurrent discontinuities. The application of the results obtained is exemplified by the propagation of a shock wave in a duct of variable cross section.

STATEMENT OF THE PROBLEM

We consider the accelerated motion of a nonstationary shock wave (SW) in a one-dimensional vortex nonisobaric flow of perfect gas. In terms of the Lagrange variables, the set of equations for this flow has the form [6]

$$\frac{\partial \ln\rho}{\partial\tau} + \frac{\gamma^2 p x^\delta}{a^2}\frac{\partial v}{\partial q} = -\frac{\delta\gamma v}{x},$$

$$\frac{\partial v}{\partial\tau} + p x^\delta \frac{\ln p}{\partial q} = 0, \qquad (1)$$

$$\frac{\partial S}{\partial\tau} = 0.$$

Here, ρ and a are the pressure and the sound velocity in the flow, respectively; v is the velocity of the flow; S is the entropy, which is related to p and a as

$$S = 2c_p\left(\ln a + \frac{\gamma-1}{2\gamma}\ln p\right) + \text{const}, \qquad (2)$$

where γ is the adiabatic exponent; q and τ are Lagrangean coordinates; and $\delta = 0$, 1, and 2 for planar, axisymmetric, and spherically symmetric flows, respectively.

The Eulerian coordinate $x = x(q, t)$ is considered as a solution of the differential equation $\partial x/\partial t = v(q, t)$. Introducing the vector $u = [\ln p, v, S]$, we come to the set of equations in the matrix form

$$\frac{\partial u}{\partial\tau} + A\frac{\partial u}{\partial q} = b, \qquad (3)$$

where the matrix A has a rank of 2 and $A[1\ldots2, 3] = 0[1\ldots2, 3]$. This set can be represented in the characteristic form [6]:

$$L^{(k)}U + \lambda_k L^{(k)}V = L^{(k)}b; \quad k = 1, \ldots, 3. \qquad (4)$$

Here, $L^{(1)} = [1, -\gamma/a, 0]$, $L^{(2)} = [1, \gamma/a, 0]$, $L^{(3)} = [0, 0, 1]$,

$$U = \left(\frac{\partial \ln p}{\partial\tau}, \frac{\partial v}{\partial\tau}, \frac{\partial S}{\partial\tau}\right), \quad V = \left(\frac{\partial \ln p}{\partial q}, \frac{\partial v}{\partial q}, \frac{\partial S}{\partial q}\right),$$

$$\lambda_{1,2} = \mp\frac{\gamma p x^\delta}{a}, \quad \lambda_3 = 0, \quad b = \left(-\frac{\delta\gamma v}{x}, 0, 0\right).$$

In the fixed coordinate system, the discontinuity $[f] = f_2 - f_1$ of the gasdynamic variables $f \in \{\ln p, v,$

1063-7842/02/4701-0018$22.00

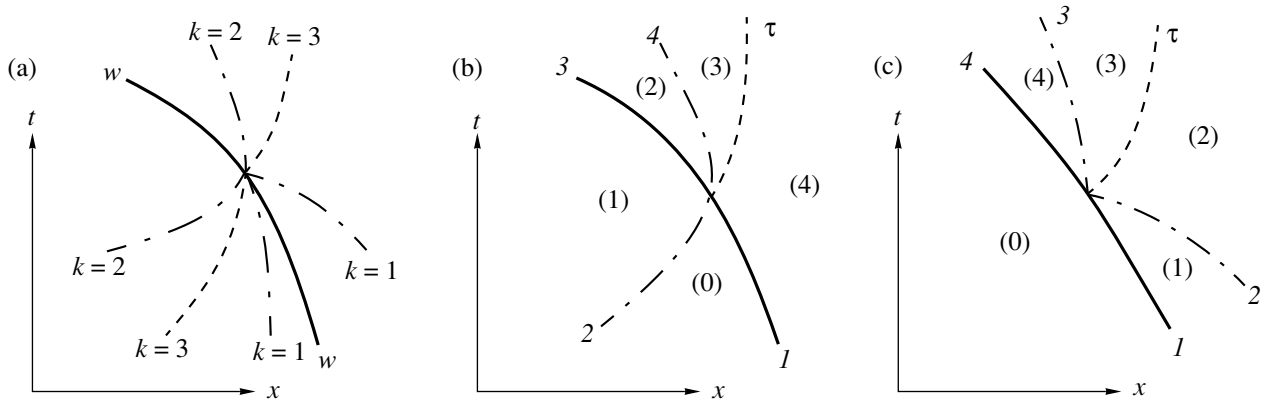


Fig. 1. Interaction of a shock wave with weak discontinuities. Parenthesized figures indicate the continuity regions of the derivatives.

$\ln a, S\}$ in an SW and the SW velocity $D = dx/dt$ relate by the Hugoniot relationships

$$\Lambda = \ln p_2 - \ln p_1 = \ln\left[(1+\varepsilon)\left(\frac{D-v_1}{a_1}\right)^2 - \varepsilon\right],$$

$$\varepsilon = \frac{\gamma-1}{\gamma+1},$$

$$v_2 - v_1 = (D-v_1)\frac{(1-\varepsilon)(J-1)}{(J+\varepsilon)}, \quad p = \frac{p_2}{p_1}, \quad (5)$$

$$\frac{\ln a_2 - \ln a_1}{2} = \ln\frac{J(1+\varepsilon J)}{(J+\varepsilon)},$$

$$\frac{\ln S_2 - \ln S_1}{c_p} = \ln\frac{J(1+\varepsilon J)J^{1/\gamma}}{(J+\varepsilon)}.$$

In passing over a shock, not only the gasdynamic variables but also their variables, $\partial f/\partial\tau$ and $\partial f/\partial q$, have discontinuities. The basic goal of the first part of the article is to express the derivatives of the main gasdynamic variables behind the shock through the main gasdynamic variables and shock parameters in front of it. The second part of the particle deals with the interaction of an SW with a weak discontinuity. As has been shown (see, e.g., [6]), the line of a weak discontinuity is necessarily coincident with one of the characteristics of set (4); in other words, there exists $k \in (1, \ldots, 3)$ such that $dq/d\tau = \lambda_k$. Hereafter, such a line will be referred to as the line of the weak discontinuity of the subscript k.

For definiteness, we assume that the SW direction coincides with that of the characteristic of the first family (Fig. 1a). Then, the SW may interfere with a weak counter discontinuity of the subscript k ($k = 1, \ldots, 3$) or with a weak cocurrent discontinuity of subscript 1. As a result, a weak discontinuity of subscript 2 and that of subscript 3 arise (reflected weak discontinuities of $k = 2$ and 3; Fig. 1a). Moreover, the SW acceleration W changes stepwise. The problem of strong–weak discontinuity interaction is stated as follows: given relationships (5) for a strong discontinuity and the discontinuities of the derivatives $\partial f/\partial\tau$ and $\partial f/\partial q$ for a counter weak discontinuity, it is necessary to find the jump in the derivatives of weak discontinuities outgoing from the point of interaction.

RELATIONSHIPS BETWEEN THE DERIVATIVES BEHIND A SHOCK WAVE AND THOSE ALONG THE SHOCK PATH

Let $w(\tau)$ be the SW path. The derivative of the gasdynamic function f_2 behind the SW with respect to τ in the direction of $w(\tau)$, the SW velocity D, and the derivatives $\partial f_2/\partial\tau_2$ and $\partial f_2/\partial q_2$ are related as follows:

$$\frac{df_2}{d\tau} = \frac{\partial f_2}{\partial\tau_2} + (D-v_2)\frac{\gamma p_2 x^\delta}{a_2^2}\frac{\partial f_2}{\partial q_2}. \quad (6)$$

As was noted, the rank of the matrix of set (3) equals two; that is, the third row of the matrix is the linear combination of the first two. In addition, the first two rows of the set involve only the derivatives of $\ln p$ and v, while the third one contains only the derivative of the function S. This means that, from (1), we can separate the subset

$$\frac{\partial\ln p}{\partial\tau} + \frac{\gamma^2 p x^\delta}{a^2}\frac{\partial v}{\partial q} = -\frac{\delta\gamma v}{x},$$
$$\frac{\partial v}{\partial\tau} + p x^\delta\frac{\partial\ln p}{\partial q} = 0, \quad (7)$$

as well as the equation

$$\frac{\partial S}{\partial s} = 0. \quad (8)$$

Then, using formulas (5), we can separately find a relation between the derivatives of $\ln p$ and v, as well as between the derivatives of S.

Let us express the derivatives of $\ln p$ and v behind the SW through the derivatives $\partial f_2/\partial\tau$ along the SW

path. To do this, we substitute $p = p_2$ and $v = v_2$ into (7). In view of relationships (6) written for $f_2 = \ln p_2$ and $f_2 = v_2$, we have the linear system for the derivatives of $\ln p$ and v behind the SW. Solving this system, we obtain

$$\frac{\partial \ln p_2}{\partial q_2} = \frac{1}{z}\left[\frac{1}{\gamma}\frac{d\ln p_2}{\partial \tau} + \frac{1}{[D-v_2]}\frac{dv_2}{\partial \tau} + \frac{\delta v_2}{x}\right] \times\left(-\frac{[D-v_2]}{p_2 x^{\delta}}\right),$$

$$\frac{\partial v_2}{\partial q_2} = \frac{1}{z}\left[\frac{1}{\gamma}\frac{d\ln p_2}{\partial \tau} + \frac{[D-v_2]}{a_2^2}\frac{dv_2}{d\tau} + \frac{\delta v_2}{x}\right]\left(-\frac{a_2^2}{\gamma p_2 x^{\delta}}\right),$$

$$\frac{\partial \ln p_2}{\partial \tau_2} = \frac{1}{z}\left[\frac{1}{\gamma}\frac{a_2^2}{[D-v_2]}\frac{d\ln p_2}{d\tau} + \frac{dv_2}{d\tau} + [D-v_2]\frac{\delta v_2}{x}\right] \times\left(\frac{\gamma[D-v_2]}{a_2^2}\right), \quad (9)$$

$$\frac{\partial v_2}{\partial \tau_2} = \frac{1}{z}\left[\frac{1}{\gamma}[D-v_2]\frac{d\ln p_2}{d\tau} + \frac{dv_2}{d\tau} + [D-v_2]\frac{\delta v_2}{x}\right],$$

$$z = 1 - \frac{[D-v_2]^2}{a_2^2} = \frac{J-1}{J(1+\varepsilon)}.$$

Using relationships (5) for the discontinuities, we can express the derivatives $df_2/d\tau$ through the derivatives $df_1/d\tau$ of the gasdynamic variables before the SW:

$$\frac{d\ln p_2}{d\tau} = \frac{d\ln p_1}{d\tau} + 2\frac{(J+\varepsilon)}{J}\left[\frac{1}{D-v_1}\frac{d(D-v_1)}{d\tau} - \frac{d\ln a_1}{d\tau}\right],$$

$$\frac{dv_2}{d\tau} = \frac{dv_1}{d\tau} + \frac{(1-\varepsilon)}{(J+\varepsilon)}\frac{d(D-v_1)}{d\tau}[(J+\varepsilon)+(1+\varepsilon)] - 2(D-v_1)\frac{(1-\varepsilon^2)}{(J+\varepsilon)}\frac{d\ln a_1}{d\tau}.$$

Introducing the designations

$$N_p = \frac{d\ln p_1}{d\tau}, \quad N_u = \frac{dv_1}{d\tau},$$
$$N_a = \frac{d\ln a_1}{d\tau}, \quad N_\delta = \frac{\delta}{x}, \quad N_D = \frac{dD}{d\tau}$$

and taking into account the expressions for the derivatives along the SW, we easily come to the desired relationships for $\ln p$ and v:

$$T_i^{(2)} = \frac{-1}{z}[\psi_p^{(i)}N_p + \psi_v^{(i)}N_v + \psi_a^{(i)}N_a + \psi_\delta^{(i)}N_\delta + \psi_D^{(i)}N_D], \quad (10)$$
$$N_i^{(2)} = \frac{1}{z}[\varphi_p^{(i)}N_p + \varphi_v^{(i)}N_v + \varphi_a^{(i)}N_a + \varphi_\delta^{(i)}N_\delta + \varphi_D^{(i)}N_D],$$

where $i = 1$ and 2,

$$T_1^{(2)} = \frac{\partial \ln p_2}{\partial q_2}, \quad T_2^{(2)} = \frac{\partial v_2}{\partial q_2}, \quad N_1^{(2)} = \frac{\partial \ln p_2}{\partial \tau_2}, \quad N_2^{(2)} = \frac{\partial v_2}{\partial \tau_2},$$

$$\psi_p^{(1)} = \tilde{\psi}_v^{(2)} = \frac{D-v_2}{\gamma p_2 x^{\delta}}, \quad \tilde{\psi}_v^{(1)} = \frac{1}{p_2 x^{d}}, \quad \psi_2^{(2)} = \frac{a_2^2}{\gamma^2 p_2 x^{d}},$$

$$\varphi_p^{(1)} = \tilde{\varphi}_v^{(2)} = 1, \quad \tilde{\varphi}_v^{(1)} = \frac{(D-v_2)\gamma}{a_2^2}, \quad \varphi_p^{(2)} = \frac{D-v_2}{\gamma},$$

$$\psi_a^{(i)} = -d_1(D-v_1)(\psi_p^{(i)} + \tilde{\psi}_v^{(i)}g_2),$$
$$\varphi_a^{(i)} = -d_1(D-v_1)(\varphi_p^{(i)} + \tilde{\varphi}_v^{(i)}g_2),$$
$$\psi_D^{(i)} = d_1(\psi_p^{(i)} + g_2\tilde{\psi}_v^{(i)}) + \tilde{\psi}_v^{(i)}d_2,$$
$$\varphi_D^{(i)} = d_1(\varphi_p^{(i)} + g_2\tilde{\varphi}_v^{(i)}) + \tilde{\varphi}_v^{(i)}d_2,$$
$$\psi_\delta^{(i)} = \psi_p^{(i)}v_2\gamma, \quad \varphi_\delta^{(1)} = \varphi_p^{(1)}v_2\gamma,$$
$$\varphi_\delta^{(2)} = (\varphi_p^{(2)} - 1)v_2\gamma,$$
$$\psi_v^{(i)} = \tilde{\psi}_v^{(i)} - \psi_D^{(i)}, \quad \varphi_v^{(i)} = \tilde{\varphi}_v^{(i)} - \varphi_D^{(i)}.$$

Here,

$$d_1 = \frac{\partial \Lambda}{\partial D} = \frac{2(J+\varepsilon)}{J(D-v_1)},$$
$$d_2 = \frac{\partial [v]}{\partial D} = \frac{(1-\varepsilon)(J-1)}{(J+\varepsilon)},$$
$$g_2 = \frac{\partial [v]}{\partial \Lambda} = (D-v_1)\frac{(1-\varepsilon^2)J}{(J+\varepsilon)^2}.$$

In practice, it is sometimes convenient to replace N_a by the function $N_S = dS/d\tau$, which characterizes the vorticity of the flow. N_S, N_a, and N_p are related as [2]

$$N_a = \frac{d\ln a}{d\tau} = \frac{\gamma-1}{2\gamma}\frac{d\ln p}{d\tau} + \frac{1}{2c_p}\frac{dS}{d\tau} = \frac{\gamma-1}{2\gamma}N_p + \frac{1}{2c_p}N_S.$$

It is clear that the coefficients before N_S differ from the associated coefficients before N_a by a factor of $1/(2c_p)$,

$$\tilde{\psi}_S^{(i)} = \psi_a^{(i)}/(2c_p), \quad \tilde{\varphi}_S^{(i)} = \varphi_a^{(i)}/(2c_p),$$

and the new and old coefficients before N_p are related as

$$\tilde{\psi}_p^{(i)} = \psi_p^{(i)} - \frac{\gamma-1}{2\gamma}\psi_a^{(i)}, \quad \tilde{\varphi}_p^{(i)} = \varphi_p^{(i)} - \frac{\gamma-1}{2\gamma}\varphi_a^{(i)}.$$

To set a relationship between the derivatives of the function S, we note that the derivatives $\partial S/\partial \tau$ vanish on both sides of the discontinuity by virtue of (8). Hence,

$$N_3^{(2)} = \frac{\partial \ln S_2}{\partial q_2} = \frac{a_2^2}{\gamma p_2 x^{d}(D-v_2)}\left[\frac{d\ln S}{d\tau} + \frac{d[S]}{d\tau}\right].$$

Using the last relationship in (5), we rearrange the derivative on the right-hand side to obtain (in terms of the above designations)

$$N_3^{(2)} = \frac{a_2^2}{\gamma p_2 x^d (D-v_2)}\left[N_S - \frac{2c_p(1-\varepsilon)\varepsilon(J-1)^2}{(1+\varepsilon)J(1+\varepsilon J)} \times\left(N_a + \frac{1}{D-v_1}(N_v - N_D)\right)\right]. \quad (11)$$

RELATIONSHIPS BETWEEN THE DERIVATIVES BEHIND THE SHOCK AND BASIC FLOW NONUNIFORMITIES BEFORE THE SHOCK

Now let us express the derivatives behind the SW through the discontinuity acceleration N_D, as well as through the functions

$$N_1 = \frac{\partial \ln p_1}{\partial \tau_1}, \quad N_2 = \frac{\partial v_1}{\partial \tau_1}, \quad N_3 = \frac{\partial S_1}{\partial q_1},$$

which are, respectively, flow nonisobaricity, flow acceleration, and flow vorticity (so-called flow nonuniformities) before the SW.

The derivative $df_1/d\tau$ of the gasdynamic function f_1 before the SW is related to the derivatives $df_1/d\tau_1$ and df_1/dq_1 as

$$\frac{df_1}{d\tau} = \frac{\partial f_1}{\partial \tau_1} + (D-v_1)\frac{\gamma p_1 x^\delta}{a_1^2}\frac{\partial f_1}{\partial q_1}. \quad (12)$$

Using set (1) written for $f = f_1$, as well as relationships (12), one can easily express N_p, N_v, and N_S through the basic flow nonuniformities:

$$N_p = N_1 - (D-v_1)\frac{\gamma}{a_1^2}N_2, \quad N_S = (D-v_1)\frac{\gamma p_1 x^\delta}{a_1^2}N_3,$$

$$N_v = N_2 = \frac{1}{\gamma}(D-v_1)[N_1 + \gamma v_1 N_\delta].$$

Substituting these expressions into formulas (10) and (11) in place of N_p, N_v, and N_S and designating $N_4 = N_\delta$ and $N_5 = N_D$ yields the desired relation between the derivatives behind the SW and the basic flow nonuniformities before it:

$$N_1^{(2)} = \frac{\Gamma(a_2)}{z}\sum_{k=1}^{5} A_{1k}N_k, \quad T_1^{(2)} = \frac{-1}{z p_2 x^\delta}\sum_{k=1}^{5} A_{2k}N_k,$$

$$N_2^{(2)} = \frac{1}{z}\sum_{k=1}^{5} A_{2k}N_k,$$

$$T_2^{(2)} = \frac{-a_2}{z p_2 x^\delta \gamma}\sum_{k=1}^{5} A_{1k}N_k - \frac{a_2^2 v_2}{p_2 x^\delta}N_4, \quad (13)$$

$$N_3^{(2)} = N_3 - \sigma(J-1)^2\sum_{k=1}^{5} A_{3k}N_k.$$

Here,

$$A_{15} = d_1 f_1 + d_2\tilde{s}_2, \quad A_{25} = d_1 f_2 + d_2\tilde{c}_2,$$

$$A_{35} = f_3/(D-v_1),$$

$$f_1 = g_1\tilde{c}_2 + g_2\tilde{s}_2, \quad f_2 = g_1\tilde{s}_2 + g_2\tilde{c}_2, \quad f_3 = 1,$$

$$g_1 = \Gamma^{-1}(a_2), \quad \Gamma(a) = \frac{\gamma}{a}, \quad \tilde{c}_2 = 1, \quad \tilde{s}_2 = \frac{D-u_2}{a_2},$$

$$A_{14} = A_{15}v_1(D-v_1) + \tilde{s}_2(v_2(D-v_2) - v_1(D-v_1)),$$

$$A_{24} = A_{25}v_1(D-v_1) + \tilde{c}_2(v_2(D-v_2) - v_1(D-v_1)),$$

$$A_{34} = A_{35}v_1(D-v_1),$$

$$A_{13} = \alpha d_1(D-v_1)f_1, \quad A_{23} = \alpha d_1(D-v_1)f_2,$$

$$A_{33} = \alpha f_3,$$

$$\alpha = -\frac{\gamma p_1 x^\delta}{2c_p a_1^2}(D-u_1), \quad \sigma = \frac{a_2^2}{\gamma^2 p_2 x^\delta (D-u_2)}\frac{2c_p\varepsilon}{J(1+\varepsilon J)},$$

$$A_{12} = -A_{15}\tilde{c}_1 + \frac{\varepsilon}{1+\varepsilon}d_1(D-v_1)f_1\Gamma(a_1)\tilde{s}_1 + \tilde{c}_1\tilde{s}_2 - \frac{\Gamma(a_1)}{\Gamma(a_2)}\tilde{s}_1\tilde{c}_2,$$

$$A_{22} = -A_{25}\tilde{c}_1 + \frac{\varepsilon}{1+\varepsilon}d_1(D-v_1)f_2\Gamma(a_1)\tilde{s}_1 + \tilde{c}_1\tilde{c}_2 - \frac{\Gamma(a_1)}{\Gamma(a_2)}\tilde{s}_1\tilde{s}_2,$$

$$A_{32} = -A_{35}\tilde{c}_1 + \frac{\varepsilon}{1+\varepsilon}f_3\Gamma(a_1)\tilde{s}_1,$$

$$\tilde{c}_1 = 1, \quad \tilde{s} = \frac{D-u_1}{a_1},$$

$$A_{11} = \frac{1}{\Gamma(a_1)}A_{15}\tilde{s}_1 - \frac{\varepsilon}{1+\varepsilon}d_1(D-v_1)f_1\tilde{c}_1 - \frac{1}{\Gamma(a_1)}\tilde{s}_1\tilde{s}_2 + \frac{1}{\Gamma(a_2)}\tilde{c}_1\tilde{c}_2,$$

$$A_{21} = \frac{1}{\Gamma(a_1)}A_{25}\tilde{s}_1 - \frac{\varepsilon}{1+\varepsilon}d_1(D-v_1)f_2\tilde{c}_1 - \frac{1}{\Gamma(a_1)}\tilde{s}_1\tilde{c}_2 + \frac{1}{\Gamma(a_2)}\tilde{c}_1\tilde{s}_2,$$

$$A_{31} = \frac{1}{\Gamma(a_1)}A_{25}\tilde{s}_1 - \frac{\varepsilon}{1+\varepsilon}f_3\tilde{c}_1.$$

RELATIONSHIPS FOR A WEAK DISCONTINUITY

Before passing to the problem of strong–weak discontinuity interaction, it would be well to establish a number of relationships concerning a weak discontinuity of the subscript k. Let $q(\tau)$ be the line of a weak discontinuity of the subscript m that is defined by the equation $dq/d\tau = \lambda_m$. From the continuity condition for the gasdynamic functions in the vicinity of a weak discontinuity, it follows that the derivatives of the vector functions $u = (\ln p, v, S)$ are equal in the direction $q(\tau)$ of the weak discontinuity:

$$U_1 + \lambda_m V_1 = U_2 + \lambda_m V_2. \tag{14}$$

Subscripts 1 and 2 refer to the derivatives on the opposite sides of the discontinuity. Since u satisfies characteristic set (4), we have

$$\begin{aligned} L^{(k)}U_1 + \lambda_k L^{(k)}V_1 &= L^{(k)}b, \\ L^{(k)}U_2 + \lambda_k L^{(k)}V_2 &= L^{(k)}b, \quad k = 1, \ldots, 3 \end{aligned} \tag{15}$$

at any point in the line of this discontinuity.

Subtracting the second expression from the first one yields

$$\begin{gathered} (L^{(k)}U_1 - L^{(k)}U_2) + \lambda_k(L^{(k)}V_1 - L^{(k)}V_2) = 0; \\ k = 1, \ldots, 3. \end{gathered} \tag{16}$$

Eliminating the difference of the derivatives with respect to τ with (14), we can eventually write

$$(\lambda_k - \lambda_m)(L^{(k)}V_1 - L^{(k)}V_2) = 0; \quad k = 1, \ldots, 3. \tag{17}$$

It follows from (17) that the equalities

$$L^{(k)}V_1 - L^{(k)}V_2 = 0, \quad k \neq m \tag{18}$$

hold at the weak discontinuity of the subscript m for any $k \neq m$.

By virtue of (16), the equalities

$$L^{(k)}U_1 - L^{(k)}U_2 = 0, \quad k \neq m \tag{19}$$

are also valid.

Introducing the designation $[f] = f_2 - f_1$, we can obtain from the last two formulas

$$\left[\frac{\partial \ln p}{\partial q} + \frac{\gamma}{a}\frac{\partial u}{\partial q}\right] = 0, \quad \left[\frac{\partial S}{\partial q}\right] = 0, \tag{20}$$

$$\left[\frac{\partial \ln p}{\partial \tau} + \frac{\gamma}{a}\frac{\partial u}{\partial \tau}\right] = 0, \quad \left[\frac{\partial S}{\partial \tau}\right] = 0 \tag{21}$$

for a weak discontinuity of the subscript $k = 1$.

Similarly, for a weak discontinuity of the subscript $k = 2$, we have

$$\left[\frac{\partial \ln p}{\partial q} + \frac{\gamma}{a}\frac{\partial u}{\partial q}\right] = 0, \quad \left[\frac{\partial S}{\partial q}\right] = 0, \tag{22}$$

$$\left[\frac{\partial \ln p}{\partial \tau} + \frac{\gamma}{a}\frac{\partial u}{\partial \tau}\right] = 0, \quad \left[\frac{\partial S}{\partial \tau}\right] = 0. \tag{23}$$

Finally, for a weak discontinuity of the subscript $k = 3$ (weak contact discontinuity), the differential conditions

$$\left[\frac{\partial \ln p}{\partial q}\right] = 0, \quad \left[\frac{\partial v}{\partial q}\right] = 0, \tag{24}$$

$$\left[\frac{\partial \ln p}{\partial \tau}\right] = 0, \quad \left[\frac{\partial v}{\partial \tau}\right] = 0 \tag{25}$$

for dynamic compatibility follow from (18) and (19).

INTERACTION OF A SHOCK WAVE WITH A COUNTER WEAK DISCONTINUITY

The above relationships, which relate the derivatives of the basic gasdynamic functions at weak and strong discontinuities, allow one to effectively solve the problem of SW–weak discontinuity interaction. In this section, we will consider the interaction of SW *1* with counter weak discontinuity *2* of the subscript k ($k = 1, \ldots, 3$) (Fig. 1b).

At the point of interaction, SW *3* with an acceleration W_3 and weak discontinuities *4* and τ of subscripts 2 and 3, respectively, originate. Let us introduce the vectors of discontinuity of the derivatives behind the SW: $[V]_w = V^{(4)} - V^{(2)}$ and $[U]_w = U^{(4)} - U^{(2)}$.

Theorem 1. If an SW whose direction coincides with that of the characteristic of the first family interacts with a counter weak discontinuity, the vectors $[V]_w$ and $[U]_w$ of derivative discontinuity behind the SW (strong discontinuity) are orthogonal to the left eigenvector $L^{(1)}$; that is,

$$L^{(1)}[V]_w = 0, \quad L^{(1)}[U]_w = 0. \tag{26}$$

Proof. As follows from Fig. 1b, the differences of the derivatives $V^{(4)}$, $U^{(4)}$ and $V^{(2)}$, $U^{(2)}$ in the regions in front of and behind the point of interaction are related to the vectors $[V]_w$ and $[U]_w$ of derivative discontinuity at originating weak discontinuities τ and 4 by the obvious expressions

$$\begin{aligned} [V]_w &= V^{(4)} - V^{(2)} = (V^{(4)} - V^{(3)}) + (V^{(3)} - V^{(2)}), \\ [U]_w &= U^{(4)} - U^{(2)} = (U^{(4)} - U^{(3)}) + (U^{(3)} - U^{(2)}). \end{aligned} \tag{27}$$

Weak discontinuities 4 and τ have subscripts 2 and 3, respectively. Multiplying by the left eigenvector $L^{(1)}$ and taking into account formulas (18) and (19), we prove the theorem.

Consequence 1. When an SW interacts with a counter weak discontinuity, the product of the left eigenvector $L^{(1)}$ by the derivative $du_2/d\tau$ of the vector function u_2 along the SW path remains unchanged and

equals

$$L^{(1)}\frac{du^4}{d\tau} = L^{(1)}\frac{du^{(2)}}{d\tau} =: L^{(1)}\frac{du_2}{d\tau} = L^{(1)}b. \qquad (28)$$

Proof. Consider the discontinuity line $w(\tau)$ for the vector function $u(x, t)$. Behind the SW, the derivative of an arbitrary gasdynamic function f_2 with respect to τ is related to the derivatives $\partial f_2/\partial\tau_2$ and $\partial f_2/\partial q_2$ through (6). We multiply the second equality in (26) by $(D - v_2)\gamma p_2 x^\delta/a_2^2$ and add the result to the first equality. In view of (6), we obtain

$$L^{(1)}\frac{du^{(4)}}{d\tau} - L^{(1)}\frac{du^{(2)}}{d\tau} = 0.$$

This equality means that, when multiplied by $L^{(1)}$ on the left, the derivative of u in the direction of the strong discontinuity does not change when the strong discontinuity interacts with any arbitrary weak discontinuity:

$$L^{(1)}\frac{du^{(4)}}{d\tau} = L^{(1)}\frac{du^{(2)}}{d\tau} = \text{const} = C. \qquad (29)$$

It only remains to find the constant on the right of (29). To do this, consider the characteristic of subscript 1 arriving at the point of interaction. Since it lies in region (4), the conditions on this characteristic have the form

$$L^{(1)}U^{(4)} + \lambda_1 L^{(1)}V^{(4)} = L^{(1)}b. \qquad (30)$$

Subtracting (30) from (29), we obtain on the left-hand side

$$\left((D - v_2)\frac{\gamma p_2 x^\delta}{a_2^2} - \lambda_1\right)(L^{(1)}V^{(4)} - L^{(1)}V^{(4)}) = 0.$$

Hence, $C = L^{(1)}b$, which is the required result.

Consequence 2. The discontinuity $[W] = W_3 - W_1$ of the SW acceleration is linearly related with the discontinuities $N_i^{(1)} - N_i^{(0)}$ $(i = 1, \ldots, 3)$ of the basic flow nonuniformities near the counter weak discontinuity of the subscript k.

Proof. Consider, for example, the second relationship in (26). It can be recast as

$$(N_1^{(2)} - N_1^{(4)}) - \Gamma(a_2)(N_2^{(2)} - N_2^{(4)}) = 0.$$

The derivatives $N_1^{(2)}$, $N_2^{(2)}$ and $N_1^{(4)}$, $N_2^{(4)}$ refer to the regions immediately behind the SW. Expressing them through the derivatives in front of the wave with (13), we obtain the equality that linearly relates the discontinuity $[W]$ of the SW acceleration to the discontinuities $[N_i]$ of the basic flow nonuniformities near the kth counter weak discontinuity:

$$(A_{15} - A_{25})[W] + \sum_{k=1}^{3}(A_{1k} - A_{2k})(N_k^{(1)} - N_k^{(0)}) = 0. \qquad (31)$$

Now let us consider specific cases of the problem.

(i) $k = 1$. As follows from (21), the function N_3 remains continuous near the counter weak discontinuity of the subscript $k = 1$. At the same time, the discontinuity $[N_1]$ of the derivative $\partial \ln p/\partial\tau$ and the discontinuity $[N_2]$ relate as

$$N_1^{(1)} - N_1^{(0)} = -\Gamma(a_1)(N_2^{(1)} - N_2^{(0)}). \qquad (32)$$

Then, (31) can be recast as

$$\frac{W_3 - W_1}{N_2^{(1)} - N_2^{(0)}} = \frac{\Gamma(a_1)(A_{11} - A_{21}) - (A_{12} - A_{22})}{A_{15} - A_{25}}.$$

(ii) $k = 2$. From (23), it follows, as before, that $[N_3] = 0$ and the functions $[N_1]$ and $[N_2]$ relate as

$$N_1^{(1)} - N_1^{(0)} = \Gamma(a_1)(N_2^{(1)} - N_2^{(0)}). \qquad (33)$$

Substituting this into (31) yields

$$\frac{W_3 - W_1}{N_2^{(1)} - N_2^{(0)}} = -\frac{\Gamma(a_1)(A_{11} - A_{21}) + (A_{12} - A_{22})}{A_{15} - A_{25}}.$$

(iii) $k = 3$. From the differential conditions for dynamic compatibility at a weak contact discontinuity [see (25)], we have $[N_1] = [N_2] = 0$; hence, formula (31) is reduced to

$$\frac{W_3 - W_1}{N_3^{(1)} - N_3^{(0)}} = -\frac{A_{13} - A_{23}}{A_{15} - A_{25}}.$$

Theorem 2. A weak discontinuity of the subscript m does not originate at the point where the SW interacts with the counter weak discontinuity of the subscript k if the vectors $[V]_w$ and $[U]_w$ of discontinuity of the derivatives behind the SW and the left eigenvector $L^{(m)}$ are orthogonal to each other; that is, if

$$L^{(m)}[V]_w = 0, \quad L^{(m)}[U]_w = 0; \quad m = 2, 3. \qquad (34)$$

Proof. We will prove the statement for $m = 2$; for $m = 3$, it is proved in a similar way. Multiplying equalities (27) by the eigenvector $L^{(2)}$ and taking into account formulas (18) and (19) yields

$$\begin{aligned} L^{(2)}[V]_w &= L^{(2)}(V^{(4)} - V^{(3)}), \\ L^{(2)}[U]_w &= L^{(2)}(U^{(4)} - U^{(3)}). \end{aligned} \qquad (35)$$

If the rights of these equalities are zero, this means that the vectors of derivative discontinuity at weak discontinuity *4* are orthogonal to all three eigenvectors. Since the latter are linearly independent, the orthogonality takes place only if $V^{(4)} - V^{(3)} = U^{(4)} - U^{(3)} = 0$, that is, if discontinuities near characteristic *4* are absent.

Let us examine the criteria that weak discontinuity *4* does not originate. The second relationship in (35) can then be recast as

$$(N_1^{(2)} - N_1^{(4)}) + \Gamma(a_2)(N_2^{(2)} - N_2^{(4)}) = 0.$$

Expressing the derivatives $N_1^{(2)}$, $N_2^{(2)}$ and $N_1^{(4)}$, $N_2^{(4)}$ through the derivatives in front of the SW with (13), we come to

$$(A_{15} + A_{25})[W] + \sum_{k=1}^{3} (A_{1k} + A_{2k})(N_k^{(1)} - N_k^{(0)}) = 0.$$

This relationship along with (31) forms the set of linear homogeneous equations in the variables $[W]$ and $(N_k^{(1)} - N_k^{(0)})$ $(k = 1, \ldots, 3)$ that is conveniently recast as

$$\begin{aligned} A_{15}[W] + \sum_{k=1}^{3} A_{1k}(N_k^{(1)} - N_k^{(0)}) &= 0, \\ A_{25}[W] + \sum_{k=1}^{3} A_{2k}(N_k^{(1)} - N_k^{(0)}) &= 0. \end{aligned} \tag{36}$$

Using formulas (21), (23), and (25), one can express the discontinuities $(N_k^{(1)} - N_k^{(0)})$ $(k = 1, \ldots, 3)$ through the discontinuity of one of the nonuniformities. The nontrivial solutions of the thus-obtained set of linear homogeneous equations will serve as criteria for the absence of weak reflected discontinuity *4*.

(i) $k = 1$, $m = 1$. In this case, $N_3^{(1)} = N_3^{(0)}$; hence, by virtue of (32), set (36) takes the form

$$A_{15}[W] - (\Gamma(a_1)A_{11} - A_{12})(N_2^{(1)} - N_2^{(0)}) = 0,$$
$$A_{25}[W] - (\Gamma(a_1)A_{21} - A_{22})(N_2^{(1)} - N_2^{(0)}) = 0.$$

This set has nontrivial solutions if

$$A_{15}(\Gamma(a_1)A_{21} - A_{22}) = A_{25}(\Gamma(a_1)A_{11} - A_{12}). \tag{37}$$

(ii) $k = 2$, $m = 1$. With such k, expression (33) and the equality $N_3^{(1)} - N_3^{(0)}$ are valid, so that the nontrivial solutions of set (36) are found if

$$A_{15}(\Gamma(a_1)A_{21} + A_{22}) = A_{25}(\Gamma(a_1)A_{11} + A_{12}). \tag{38}$$

With (13) it is easy to check that Eqs. (37) and (38) has the same analytic solution

$$J = 2\varepsilon^{3/2}\left(\frac{1 + \sqrt{\varepsilon}}{1 - \varepsilon}\right)^2, \tag{39}$$

which implies that the interaction without a reflected discontinuity is possible only if $\varepsilon > 1/4$, $\gamma > 5/3$, and $k = 1$.

(iii) $k = 3$, $m = 1$. Substituting the differential conditions for dynamic compatibility at a weak contact discontinuity [see (25)] into set (36), one can easily find the following criterion for the absence of a weak discontinuity:

$$A_{15}A_{23} - A_{25}A_{13} = 0.$$

With the expressions for the associated coefficients, this equality is reduced to the form

$$2(1 + \varepsilon)(J + \varepsilon)p_1 x^{\delta}/(D - v_1) = 0,$$

from which it follows that the interaction of the SW with a weak contact discontinuity without generating reflected weak discontinuity *4* is impossible.

INTERACTION OF A SHOCK WAVE WITH A COCURRENT WEAK DISCONTINUITY

As was noted above, if the direction of SW *1* coincides with that of characteristic *2* of the second family, the SW may interact with the cocurrent weak discontinuity of subscript 1 (Fig. 1c). The result of the interaction is the discontinuity $[W] = W_4 - W_1$ of the SW acceleration, as well as weak discontinuities *3* and τ of subscripts 2 and 3, respectively.

Theorem 1. If the SW interacts with the weak cocurrent discontinuity of subscript 1, the eigenvector $L^{(k)}$ $(k = 1, \ldots, 3)$ is orthogonal to the differences $[V]_w - [U]_w$ and $[V]_k - [U]_k$, where the former is the difference of the derivatives behind the SW and the latter is that near a weak discontinuity of the subscript k; that is,

$$L^{(k)}([V]_w - [V]_k) = 0,$$
$$L^{(k)}([U]_w - [U]_k) = 0; \quad k = 1, \ldots, 3.$$

Consequence 1. The discontinuity $[W] = W_4 - W_1$ of the SW acceleration is linearly related to the discontinuity of the path curvature $N_2^{(2)} - N_2^{(1)}$ at the weak cocurrent discontinuity of subscript 1; that is,

$$(A_{15} - A_{25})[W] = -2\Gamma(a_2)(N_2^{(2)} - N_2^{(1)}).$$

Consequence 2. Weak discontinuity *3* does not originate if

$$A_{15} = A_{25}. \tag{40}$$

All the three statements are proved as those in the previous section.

Using expressions (13) for the coefficients A_{ij}, one can show that equality (40) holds if

$$J = \frac{4\varepsilon^2}{1 - 3\varepsilon}.$$

Thus, the interaction without the reflected discontinuity is possible if $\varepsilon > 1/4$ and $\gamma > 5/3$.

THE CHESTER–WHITHAM FORMULA

Let us turn back to the interaction of an SW with a counter weak discontinuity. For one specific case of great importance, formula (28) is likely to be first derived by Whitham [7, 8]. He analyzed the results reported in [9–11], where a shock wave propagated in a stationary gas through a duct with a small cross-sec-

tional discontinuity. This problem is a specific case of the more general problem of discontinuity breakdown in a variable-section duct [6]. In essence, the general problem involves two subproblems: the propagation of a shock wave in a constant-section duct and the flow of a gas in a variable-section duct.

In [9], the relation between a small variation of the relative velocity $M = D/a$ of the shock and the variation of the cross-sectional area A of the duct was derived based on the linearization of the relationships at the cross-sectional discontinuity:

$$d\ln A = f(M)dM.$$

Whitham noticed that the same result can be obtained if one writes the condition for the characteristic of the second family in the flow behind the shock and, instead of p_2, u_2, and a_2, substitutes their associated expressions (in terms of M) for the shock wave [see (5)]. If the wave propagates in a stationary gas with the parameters p_1, u_1, and a_1, we have

$$\frac{p_2}{p_1} = (1+\varepsilon)M^2 - \varepsilon, \quad \frac{u_2}{a_1} = (1-\varepsilon)\left(M - \frac{1}{M}\right),$$

$$\frac{a_2}{a_1} = \sqrt{\left[(1+\varepsilon) - \frac{\varepsilon}{M^2}\right][(1-\varepsilon) + \varepsilon M^2]}.$$

Whitham called such an expedient the rule of characteristics and assumed that it can also apply to other cases [7]. Formula (28) proved in our work generalizes this rule for the interaction of a shock wave propagating in a vortex nonisobaric one-dimensional flow with a counter discontinuity of an arbitrary subscript. Also, this formula allows one to derive approximate analytical solutions for the interaction of a shock wave with a Riemann wave, a shear layer, etc.

ACKNOWLEDGMENTS

The author thanks É.A Tropp and V.N. Uskov for the valuable discussion.

This work was financially supported by INTAS project no. 99-785.

REFERENCES

1. C. Truesdell, J. Aeronaut. Sci. **19**, 826 (1952).
2. M. J. Lighthill, Philos. Mag. **40**, 214 (1949).
3. M. J. Lighthill, J. Fluid Mech. **2**, 1 (1957).
4. V. V. Rusanov, Preprint No. 18, Inst. Prikl. Mat. AN SSSR (Inst. of Applied Mathematics, Academy of Sciences of USSR, Moscow, 1973).
5. A. L. Adrianov, A. L. Starykh, and V. N. Uskov, Interference of Stationary Gasdynamic Discontinuities, Ed. by Ya. I. Shokin (Nauka, Novosibirsk, 1995).
6. B. L. Rozhdestvenskiĭ and N. N. Yanenko, *Systems of Quasi-linear Equations and Their Applications to Gas Dynamics* (Nauka, Moscow, 1968).
7. G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974; Mir, Moscow, 1977).
8. G. B. Whitham, J. Fluid Mech. **4**, 337 (1958).
9. B. Chester, Mekhanika, No. 6, 76 (1954).
10. R. F. Chisnell, J. Fluid Mech. **2**, 286 (1957).
11. R. F. Chisnell, Proc. R. Soc. London **232**, 350 (1955).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 26–33. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 28–35.

GAS DISCHARGES, PLASMA

Influence of the Cathode and Anode Jets on the Properties of a High-Current Electric Arc

A. A. Bogomaz, A. V. Budin, V. A. Kolikov, M. E. Pinchuk, A. A. Pozubenkov, and F. G. Rutberg

Institute of Problems of Electrophysics, Russian Academy of Sciences, St. Petersburg, 191065 Russia
e-mail: rutberg@iperas.spb.su

Received May 8, 2001

Abstract—A study is made of the effects related to the formation of electrode jets in discharges in hydrogen and air at a current of 10^5–10^6 A, a current growth rate of 10^{10} A/s, an initial pressure of 0.1–4.0 MPa, and a discharge gap length of 5–40 mm. After secondary breakdown, jets are observed in a semitransparent discharge channel expanding with a velocity of $(4–7) \times 10^2$ m/s. The formation of shock waves in the interaction of the jets with the ambient gas and the opposite electrode is observed by the shadowgraphy method. Seventy microseconds after the beginning of the discharge, the pressure of the metal vapor plasma near the end of the tungsten cathode amounts to 177 MPa. The brightness temperature in this case is $T = 59 \times 10^3$ K, the average ion charge number is $\bar{m} = 3.1$, and the metal vapor density is $n = 5.3 \times 10^{19}$ cm^{-3}. After 90 μs, the average ion charge number and the metal vapor density near the anode end are $\bar{m} = 2.6$ and $n = 7.4 \times 10^{19}$ cm^{-3}, respectively. Based on the experimental data, possible reasons for the abnormally high values of the total voltage drop near the electrodes (up to ~1 kV) are discussed. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Electric arcs with currents of ≥1 MA are used in a number of electrophysical devices, such as electric-discharge accelerators, current switches, and pulsed generators of high-enthalpy plasma. In this case, the power deposited in the arc channel attains 10^8–10^{10} W and most of the power is released near the electrodes. This, in turn, results in the formation of electrode jets. The cathode and anode jets produced in arcs with currents up to several tens of kiloamperes have long been known [1, 2] and have been studied in sufficient detail. The velocities of the cathode and anode jets, the temperature and density of charged particles near the origins of these jets, the electrode erosion rate, etc. were determined. The results of these studies are reviewed in [3–8]. In a number of papers, the effect of the interaction of the cathode and anode jets with each other and with the opposite electrode was studied (see, e.g., [9–11]). The peak current in these studies usually did not exceed 10 kA. In megaampere arcs, at a high level of energy deposition and high pressures of the working gas, the electrode jets significantly affect both the properties of the arc itself and its heat exchange with the ambient gas. In this case, the electrode erosion rate attains 10^{-3}–10^{-2} g/C [12, 13]. At a peak current of 60 kA, the total voltage drop near the electrodes amounts to ~500 V [14]. In megaampere arcs, the electrode voltage drop attains ~1 kV [15–18].

As was noted in [19], quasi-steady discharges with a current density of 10^5–10^7 A/cm^2 (which is typical of the above discharge conditions) possess specific features. One of them is the sharp increase in the potential drop (up to several hundreds of volts) near the anode and the cathode. Explanation of this effect is still lacking. In this paper, we discuss probable reasons for the abnormally high values of the voltage drop. It is examined how the jets in megaampere arcs influence the arc properties and the heat exchange with the ambient gas, because both this influence and arcs of this type themselves are still poorly understood.

DISCHARGE CHAMBER AND DIAGNOSTIC METHODS

The initial design of the discharge chamber was similar to that of high-power electric-discharge accelerators [20]. When modeling the operation of such accelerators, the chamber case served as an anode. An initiating wire was placed in a narrow gap between a conical cathode and the chamber wall. To obtain an axially symmetric discharge, the design of the chamber was changed according to Fig. 1.

High-current arcs in hydrogen and air were studied at an initial pressure of 0.1–4.0 MPa, a peak current of $(0.5–5.0) \times 10^5$ A, and a current growth rate of $(0.5–1.0) \times 10^{10}$ A/s. The distance between tungsten electrodes 10 mm in diameter was varied from 5 to 40 mm.

1063-7842/02/4701-0026$22.00 © 2002 MAIK "Nauka/Interperiodica"

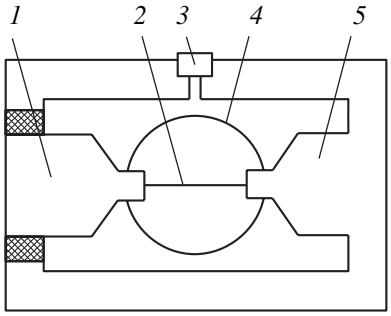


Fig. 1. Schematic of the diagnostic discharge chamber: (*1*) cathode, (*2*) initiating wire, (*3*) pressure gauge, (*4*) diagnostic windows, and (*5*) anode.

The discharge was initiated by exploding a copper or stainless-steel wire 0.15 mm in diameter.

OPTICAL MEASUREMENTS

A schematic of the arc shadowgraphy measurements is shown in Fig. 2. The arc radiation was attenuated with the help of interference filters.

The brightness temperature of different segments of the arc was evaluated from the blackening density of the high-speed photographs of the arc. The arc was photographed through different filters (Fig. 3). The blackening density from a Podmoshensky capillary source with a brightness temperature of 40×10^3 K was taken as a reference blackening density. A schematic of measurements is shown in Fig. 3.

The high-speed spectrography of the discharge was performed through neutral filters with an optical thickness of 0.5–2. Such an optical thickness range corresponds to a 100-fold variation in the image source intensity.

RESULTS

The waveforms of the current and voltage across the discharge gap are shown in Fig. 4. After the explosion of the initiating wire, even before secondary breakdown, we can see small conical jets flowing from the anode and cathode. The jets, which are seen against the background of low-intensity, chaotically expanding fragments of the initiating wire, effuse from the central parts of the electrodes.

Using neutral filters with a high optical density, we observe the expansion of a semitransparent discharge channel after secondary breakdown. We can see how, at first, the cathode and, then, the anode jet form inside the channel. The jets effuse from the large area of the ends of tungsten electrodes 6 mm in diameter (Fig. 5).

The semitransparent channel of the discharge in hydrogen expands almost symmetrically with a velocity of $(2–7) \times 10^2$ m/s, depending on the initial pressure (Fig. 6), which agrees with our previous results [17].

For air at atmospheric pressure, the expansion velocity is 4.1×10^2 m/s at a peak current of 100 kA. At a tungsten electrode diameter of 6 mm, a high-power cathode jet is formed 70 μs after the beginning of the discharge. As the discharge current increases, the current density increases to 5×10^5–10^6 A/cm^2, whereas the diameter of the discharge channel decreases. The jet diameter decreases because the magnetic field pressure increases with the current and, by this time, reaches a level of 177 MPa. In this case, the maximum brightness temperature in the center of the jet near its origin, at a distance of several millimeters from the cathode, reaches 59×10^3 K. The brightness temperature near the anode reaches its maximum value by 90 μs and amounts to 52×10^3 K.

At instants corresponding to the maximum discharge current (Fig. 4), the plasma becomes opaque, the entire surface of the discharge channel glows, and no jets are observed. Later, in the experiments with copper electrodes 10 mm in diameter, we observe a compres-

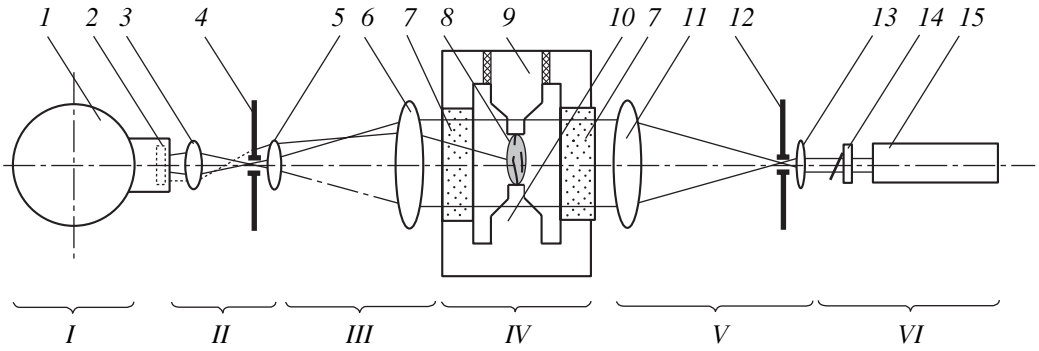


Fig. 2. Schematic of the shadowgraphy measurements: (*I*, *1*) ZhLV-2 high-speed camera, (*II*) matching objective, (*III*) receiving part, (*IV*) discharge chamber, (*V*) collimating part, (*VI*) illumination source; (*2*) light filters, (*3*) second component of the matching objective, (*4*) visualizing diaphragm, (*5*) first component of the matching objective, (*6*) receiving lens, (*7*) diagnostic windows, (*8*) arc, (*9*) cathode, (*10*) anode, (*11*) collimating lens, (*12*) beam cleaner, (*13*) condenser, (*14*) electrodynamic shutter, and (*15*) LGN-402 argon laser.

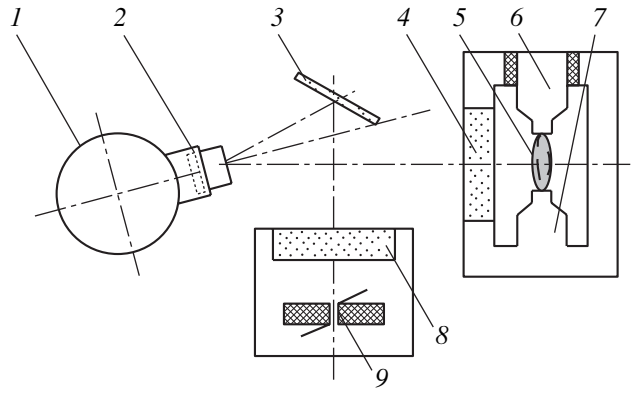


Fig. 3. Schematic of the brightness temperature measurements: (*1*) ZhLV-2 high-speed camera; (*2*) neutral filters, broadband green filter, and 5500 Å interference filter; (*3*) mirror; (*4*, *8*) windows; (*5*) arc; (*6*) cathode; (*7*) anode; and (*9*) reference capillary source.

sion shock. It corresponds to the front of a shock wave arising when the high-power cathode jet interacts with the ambient gas. The compression shock is located far behind the anode (Fig. 7).

The formation of cathode and anode jets is seen more clearly in air discharges at atmospheric pressure, a tungsten electrode diameter of 6 mm, and a peak current of up to 100 kA. In this case, due to a higher transparency of the discharge, the expansion of the channel and the subsequent formation of jets can be observed on the shadow photographs over a longer period of time (Fig. 8).

With tungsten electrodes 6 mm in diameter, a high-power anode jet in an air discharge is formed in ~100 μs. As the intensity of the anode jet increases, a brighter region associated with the interaction between the cathode and anode jets is displaced from the anode toward the cathode. When the anode and cathode jets have the same intensities, a luminous region is equally spaced from the cathode and the anode (Fig. 9). As the intensity of the anode jet increases, the shocked gas region is displaced toward the cathode (Fig. 10).

For a discharge in hydrogen, a nonmonotonic increase in the arc voltage and the steady-state gas pressure in the chamber was observed as the discharge gap length increased (Fig. 11). In addition, discharge voltage jumps with an amplitude of about 400 V were observed. Such an increase in the voltage coincided in time with the instant the electrode spot was displaced from the side surface to the end of the electrode.

Using the shadowgraphy technique, we observed the intense motion of gas throughout the entire discharge volume in the course of the formation of the cathode and anode jets. The direction of gas motion near the wall of the discharge chamber was opposite to the direction of the jet motion, and the gas velocity was 100–200 m/s.

DISCUSSION OF RESULTS

Depending on the duration of electrode heating, the current density, the dimensions and material of the electrodes, etc., we observed two types of jets associated with the formation of electrode spots.

The jets of the first type are associated with the formation of rapidly moving electrode spots. They effuse from the central regions of the electrodes and are formed already during the wire explosion. The jets are seen against the background of low-intensity, chaotically expanding fragments of the wire in the first microseconds after the wire explosion, even before secondary breakdown. The jets that are seen in the semitransparent discharge channel in Fig. 5 are classed as jets of the first type. Their formation can be caused by a confluence of individual jets emerging from rapidly moving spots [4].

The jets of the second type effuse from a common melt bath which is formed on the electrode upon its heating. The heating time depends on the electrode material, electrode diameter, and the current density.

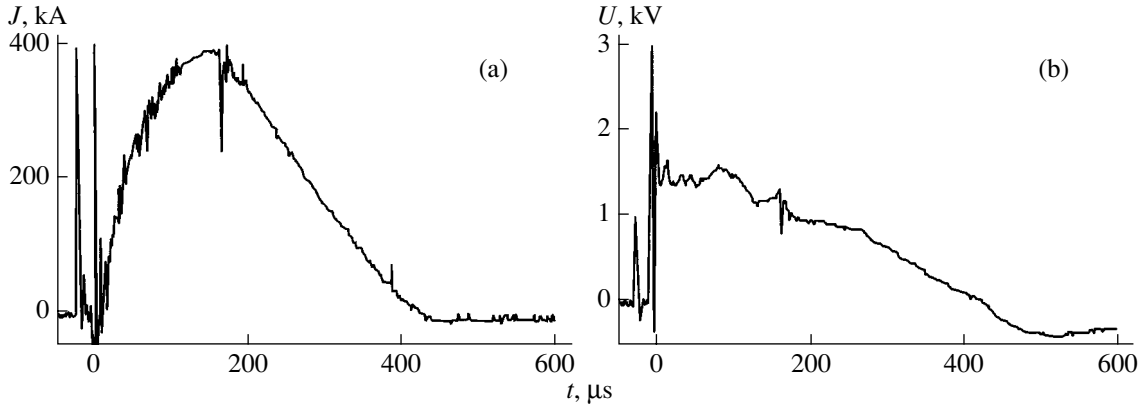


Fig. 4. Waveforms of (a) the discharge current and (b) voltage. The initial hydrogen pressure is 1 MPa, and the diameter of tungsten electrodes is 6 mm.

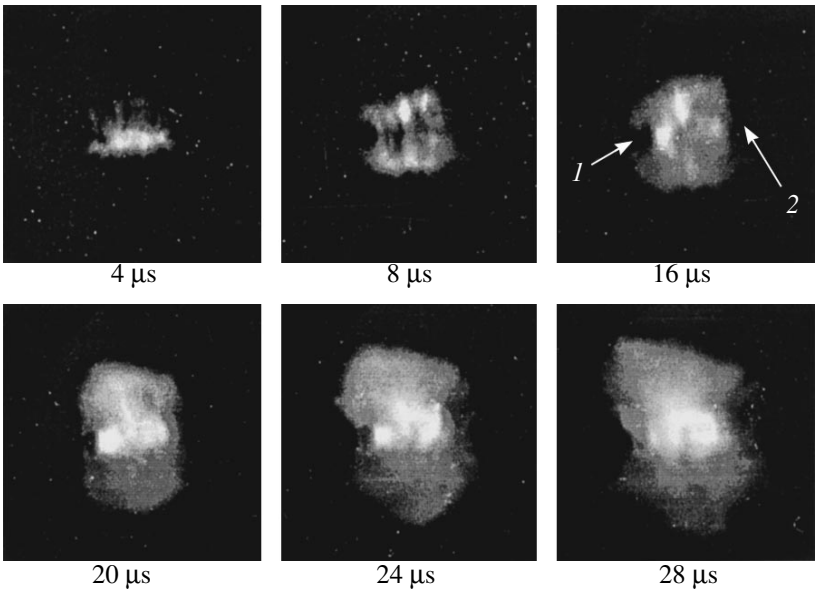


Fig. 5. Formation of the cathode and anode jets during the expansion of the semitransparent channel of a hydrogen discharge: (*1*) cathode and (*2*) anode.

Under our experimental conditions, for a tungsten electrode 6 mm in diameter at a current density of 5×10^5 A/cm^2, the heating time amounts to 70 μs. Apparently, both types of jets can exist simultaneously. On the anode (Fig. 13), a narrow and long darker central jet of the second type is seen against the brighter background of a broad conical jet of the first type.

As was stated in [4, 5], the formation of second-type jets on the anode can be related to the pinch effect. At currents above 10 kA, the magnetic field pressure becomes close to the gas-kinetic pressure. This leads to the compression of the arc in the effusion region and its additional heating.

The central jet, in its appearance and by the generation mechanism, resembles a plasma focus arising when a current-carrying shell is compressed by the magnetic field of the discharge [21].

Let us determine the parameters of the erosion plasma near the end of the cathode for the jet of the second type 70 μs after the beginning of the discharge. These parameters can be evaluated from the values of the brightness temperature T and pressure P. According to [22], the metal vapor density in the cathode spot is 10^{19}–10^{20} cm^{-3}. Consequently, we may assume that the plasma is in local thermal equilibrium and the brightness temperature corresponds to the equilibrium temperature. The metal vapor density n and the average ion charge number $\bar{m}$ were determined from the set of equations [23]

$$i\left(\bar{m} + \frac{1}{2}\right) = kT \ln \frac{AT^{3/2}}{\bar{m}n}, \quad (1)$$

$$P = n(1 + \bar{m})kT, \quad A = 6 \times 10^{21} \text{ cm}^{-3} \text{ eV}^{-3/2},$$

where i is the mean value of the ionization energy for tungsten at the point $\bar{m} + 1/2$. The value of P was estimated from the position of the density jump by the

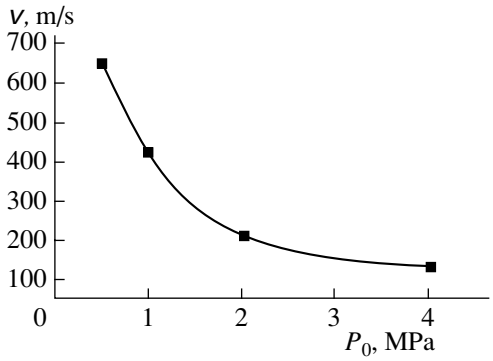


Fig. 6. Channel expansion velocity v in a hydrogen discharge vs. initial pressure P_0.

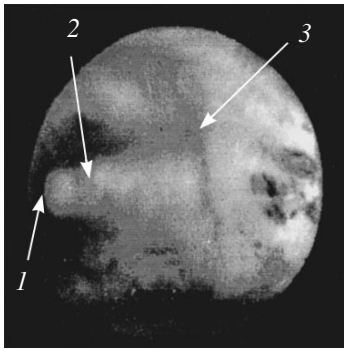


Fig. 7. Discharge in hydrogen. The formation of a shock wave generated by a cathode jet: (*1*) cathode, (*2*) anode position, and (*3*) compression shock.

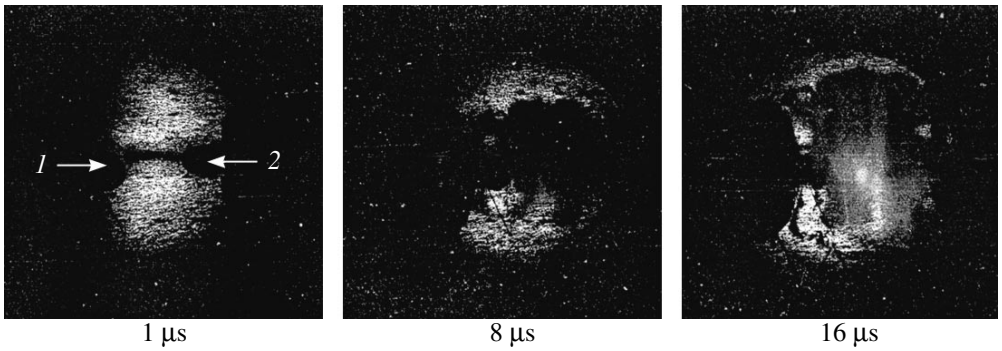


Fig. 8. Shadow photograph of an expanding discharge in air: (*1*) cathode and (*2*) anode.

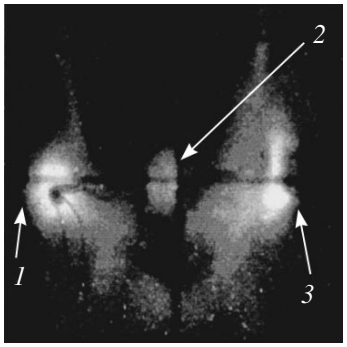


Fig. 9. Formation of a luminous region in the middle of the interelectrode gap: (*1*) cathode, (*2*) luminous region in the center of the discharge, and (*3*) anode.

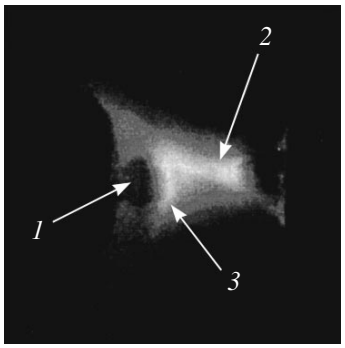


Fig. 10. Shocked region formed when the anode jet flows around the cathode: (*1*) cathode, (*2*) anode jet, and (*3*) shocked region.

formula [24]

$$P = P_H\left(\frac{l}{0.64d}\right)^2.$$

Here, P_H is the gas pressure in an unperturbed gas, l is the distance of the jump from the electrode, and d is the electrode diameter. The value of l/d was found to be to 3.6. The value of P_H was measured by a pressure gauge connected through a channel to the discharge chamber. The gauge readings corresponded to the pressure near the chamber wall. Therefore, the value P_H = 4 MPa, which corresponds to the measured value of l, is underestimated. In this case, we have P = 126 MPa, which is the lower limit of the pressure P. To determine the P value more precisely, it was assumed that the magnetic pressure is equal to the gas-kinetic pressure near the jet origin [4, 5]. This assumption is based on the observed compression of the jet. The magnetic field pressure is equal to P[MPa] = $1.6 \times 10^{-10}\ J^2/r^2$[A/cm^2], so that P = 177 MPa for $J = 3.16 \times 10^5$ A and r = 0.3 cm. Equations (1) for a tungsten plasma with $T = 59 \times 10^3$ K and P = 177 MPa have the solution $\bar{m}$ = 3.1 and $n = 5.3 \times 10^{19}$ cm^{-3}.

By the instant t = 90 μs, the brightness temperature near the anode is 52×10^3 K and the pressure is P = 190 MPa. The average ion charge number $\bar{m}$ and the ion density n of the anode plasma, determined by the same method as that used near the cathode, are $\bar{m}$ = 2.6 and $n = 7.4 \times 10^{19}$ cm^{-3}, respectively. The higher values of n at the anode correspond to a higher rate of anode erosion if we assume that, as in [2], the jets emerging from the anode and the cathode have the same velocities. This agrees with the results obtained in [12].

The values of $\bar{m}$ and n determined near the cathode and the anode from the measured values of T and P with the help of Eqs. (1) are close to the values of $\bar{m}$ and n calculated in [25] for aluminum, titanium, and iron vapors. For this reason, the averaged calculated values are used for rough estimates of the conductivity and the Rosseland radiation length.

For the values of T, $\bar{m}$, and n measured near the cathode, the plasma conductivity amounts to ~400 $(\Omega\ \text{cm})^{-1}$. For a current density of $j \sim 10^6$ A/cm^2, this corresponds to a field strength of $E \sim 2.5 \times 10^3$ V/cm. The current density $j \sim 10^6$ A/cm^2 corresponds to the optical data, which show that, at the time 70–90 μs, the diameter of the electrode luminous regions that is used for estimates is close to the electrode diameter. The total voltage drop in the electrode regions amounts to ~1 kV

[15–18]. Then, the total length of the voltage drop region is $l = U/E = 1000/2500 = 0.4$ cm, the overall length of the discharge gap being 1.7 cm.

Hence, the regions in which most of the energy is released are located near the cathode and the anode, the lengths of these regions being about 0.2 cm. The Rosseland radiation length in each of these regions is of the same order of magnitude. For this reason, the plasma is semitransparent at wavelengths at which most of energy is emitted. For a fully transparent plasma with these parameters, the total recombination radiation power from 0.2-cm-long regions 6 mm in diameter will be several orders of magnitude higher than the JU value. For a semitransparent plasma, this power may be comparable with the deposited power. The following facts evidence in favor of the existence of 0.2-cm-long regions with high emissivity near the electrodes: (i) some high-speed photographs exhibit bright regions 0.2–0.3 cm in length near the ends of the electrodes (Fig. 14); (ii) a voltage drop of ~1 kV is observed in a narrow gap in length 0.2–0.4 cm between the cylindrical anode and the conical cathode of the electric-discharge accelerator in the initial stage of the discharge in the region where the initiating wire was positioned [15, 20]; and (iii), as was mentioned above, the anode voltage increases by 400 V when the arc is displaced toward the anode end.

In this case, a voltage jump of $U = 400$ V near the anode at a field strength of $E = 2500$ V/cm, determined previously from the plasma conductivity and the current density, corresponds to the length l

$$l = \frac{U}{E} = \frac{400}{2500} = 0.16 \text{ cm}.$$

One of the reasons for the high electrode voltage drops may be the interaction of the jets with the surface of the opposite electrode (Fig. 10) or with each other. In the latter case, the interaction should occur near one of the electrodes. This hypothesis is supported by the dependence of the discharge voltage on the gap length as is shown in Fig. 11. We see that the arc voltage is maximum at a length of ~15 mm. This length probably corresponds to the most efficient interaction of the cathode jet with the anode (for copper electrodes 10 mm in diameter, no anode jets were observed). As the gap length increases further, the jet interaction region disappears and the voltage decreases. The increase in the voltage during the interaction of the jets is associated with the fact that the current carriers acquire a velocity component perpendicular to the discharge axis.

Following [22], we estimate both the velocity of the cathode jet u_2 after its acceleration in the region immediately behind the space-charge region and the contribution from the gas-kinetic pressure and electromagnetic force. We consider the cathode at the instant 70 μs, when the anode jet has yet no time to form. As a control cross section, labeled 1, we choose a plasma region lying at the boundary with the space-charge region and

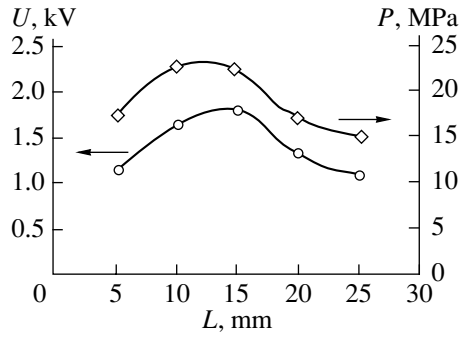


Fig. 11. Voltage drop U across the discharge gap and the steady-state pressure P in the chamber vs. interelectrode distance L for copper electrodes 10 mm in diameter. The current is ~300 kA; the initial hydrogen pressure is 1 MPa.

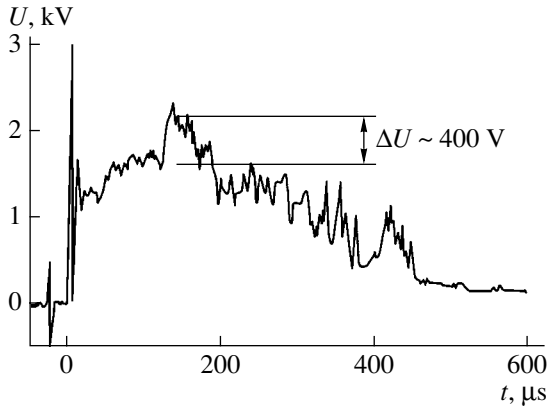


Fig. 12. Voltage jump of ~400 V observed when the electrode spot is displaced from the side surface to the end of the anode.

assume that the measured values of T and P refer precisely to this region. The second control cross section, labeled 2, lies in front of the region characterized by a sharp increase in the anode voltage. Assuming, as in [22], that $u_1 = 0$ and $P_2 \ll P_1$, we arrive at the law of conservation of momentum

$$Gu_2 = SP_1 + I, \tag{2}$$

where I is the total increase in the electromagnetic force, which, as in [22], is defined by the formula

$$I = \left(\frac{J}{c}\right)^2 \ln\frac{a_k}{a} \sim \left(\frac{J}{c}\right)^2 \ln\frac{L}{a}.$$

Here G is the total consumption of the electrode material, J is the total arc current, u_1 and u_2 are the jet velocities in the control cross sections, S is the electrode area, a_k is the arc radius in control cross section 2, a is the electrode radius, L is the discharge gap length, and c is the speed of light.

Since the jet expansion angle about the axis is ~45° and $\alpha < L$, we have $a_k \sim L$. According to [12], for the cathode made from an alloy containing 80% tungsten,

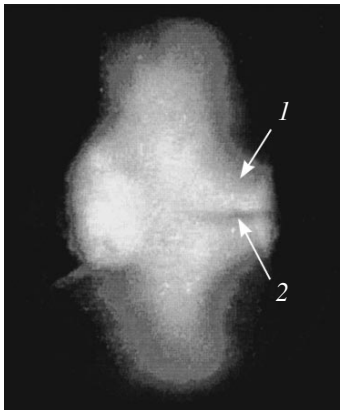


Fig. 13. (*2*) Second-type metal jet from the anode against the background of brighter radiation from (*1*) the first-type jet.

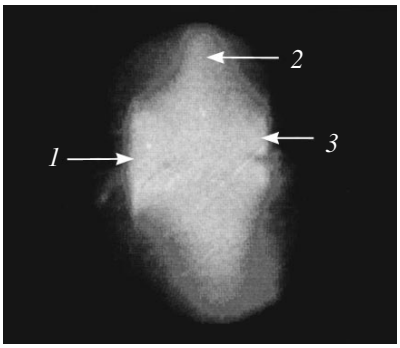


Fig. 14. (*1*, *3*) Luminous regions of length 0.2–0.3 cm near the ends of the electrodes, and (*2*) the formation of a "plasma plate" caused by the interaction between the anode and cathode jets.

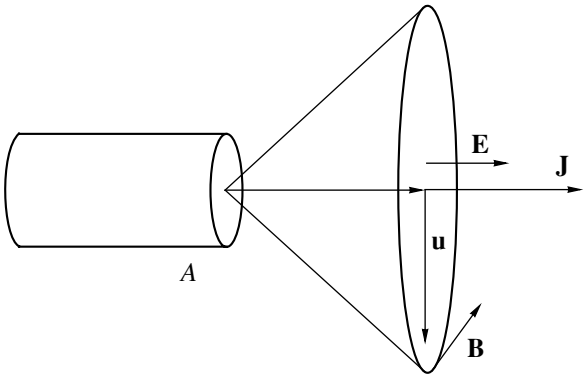


Fig. 15. Generation of an additional field **E** near the anode *A* due to the transverse velocity component **u**.

10% nickel, and 10% iron, the cathode material consumption amounts to 6.5×10^{-3} g/C. For a current of $J = 3.16 \times 10^5$ A, this corresponds to $G = 2.1 \times 10^3$ g/s. Substituting in the above formula $L = 1.7$ cm and $a = 0.3$ cm at $P_1 = 177$ MPa, we obtain $SP_1 = 4.86 \times 10^3$ N and $I = 17.3 \times 10^3$ N. Then, from formula (2), we obtain $u_2 = 1.05 \times 10^4$ m/s.

Hence, for arcs with currents $J > 3 \times 10^5$ A, a decisive factor in the acceleration of metal vapor is the action of electromagnetic forces, which are more than three times greater than gasdynamic forces in spite of a fairly high pressure near the jet origin. The estimate obtained is rather rough, because the pressure P_2 is not zero and the total cathode material consumption determined in [12] corresponds to both the vapor and drop phases. However, when estimating the velocity u_2, both errors partially compensate for each other.

One more possible cause for the high voltage drop in the electrode regions is that the directions of the current and the jet velocity do not coincide (Fig. 15).

As is well known, the additional field **E** due to the velocity **u** across the arc magnetic field **B** is equal to $\mathbf{E} = \mathbf{u} \times \mathbf{B}$, where $B = \mu J/2\pi r$. For large jet expansion angles, the transverse velocity component is comparable with the longitudinal velocity. For the latter, the literature data [1–3, 7] and the estimates performed give a value on the order of 10^4 m/s. Assuming that $J \sim 10^5$ A and $r = 3 \times 10^{-3}$ m, which correspond to the experimental data, we obtain $E \sim 10^3$ V/cm. For smaller expansion angles, the same value of E corresponds to the higher values of the discharge current.

The anode and cathode jets intensify heat exchange between the arc and the ambient gas. In our previous study, it was found that the efficiency of hydrogen heating in the discharge chamber increases as the interelectrode distance decreases [18]. At the same time, we observe an increase in the acoustic vibrations detected by the pressure gauge on the wall of the discharge chamber. Probably, such acoustic oscillations are caused by the interaction between the anode and cathode jets. The observed gas motion near the chamber wall in the direction opposite to that of the jet motion also contributes to heat exchange. Hence, high-power anode and cathode jets may be responsible for the intense vortex gas flow in the discharge chamber, thereby intensifying heat exchange.

CONCLUSIONS

(i) For a hydrogen discharge at an initial pressure of 1 MPa and a discharge current of $\sim 3 \times 10^5$ A, the parameters of the electrode erosion plasma near the cathode were found to be the following: $T = 59 \times 10^3$ K, $P = 177$ MPa, $\bar{m} = 3.1$, and $n = 5.3 \times 10^{19}$ cm^{-3}. The plasma parameters near the anode are close to these values.

(ii) The lengths of the electrode regions responsible for a total voltage drop of ~1 kV near each of the electrodes were determined to be ~0.2 cm.

(iii) High voltage drops in the electrode regions are caused by the high emissivity of the erosion plasma; the presence of the shocked regions because of the interaction of the jets with each other, the opposite electrode, and the ambient gas; and the motion of the jets across the magnetic field produced by the discharge current.

(iv) It was found that the jets produce an intense vortex gas flow in the chamber, thereby intensifying heat exchange between the arc and the ambient gas.

ACKNOWLEDGMENTS

We are grateful to A.G. Kuprin for providing us with experimental results obtained under similar conditions. This work was supported in part by the Russian Foundation for Basic Research, project no. 00-15-96605.

REFERENCES

1. A. A. Plyutto, V. N. Rozhkov, and A. T. Kapin, Zh. Éksp. Teor. Fiz. **47**, 494 (1964) [Sov. Phys. JETP **20**, 328 (1964)].
2. O. B. Bron and L. K. Sushkov, *Plasma Streams in Electric Arc of Closing Devices* (Énergiya, Leningrad, 1975).
3. G. A. Lyubimov and V. I. Rakhovskiĭ, Usp. Fiz. Nauk **125**, 665 (1978) [Sov. Phys. Usp. **21**, 693 (1978)].
4. I. A. Krinberg and E. A. Zverev, Fiz. Plazmy **25**, 88 (1999) [Plasma Phys. Rep. **25**, 82 (1999)].
5. V. A. Nemchinskiĭ, Zh. Tekh. Fiz. **52**, 35 (1982) [Sov. Phys. Tech. Phys. **27**, 20 (1982)].
6. V. A. Nemchinskiĭ, Zh. Tekh. Fiz. **53**, 235 (1983) [Sov. Phys. Tech. Phys. **28**, 146 (1983)].
7. G. A. Mesyats, *Ectons in a Vacuum Discharge: Breakdown, Spark, and Arc* (Nauka, Moscow, 2000).
8. B. Juttner, IEEE Trans. Plasma Sci. **PS-15**, 481 (1987).
9. M. A. Sultanov and L. I. Kiselevskiĭ, Teplofiz. Vys. Temp. **4**, 375 (1966).
10. G. A. Dyuzhev, V. Yu. Evlasov, and S. M. Shkol'nik, in *Proceedings of the X All-Union Conference "Low-Temperature Plasma Generators," Minsk, 1986*, Part I, p. 123.
11. A. F. Bretskikh, V. I. Sysun, and Yu. D. Khromoĭ, in *Proceedings of the XI All-Union Conference "Low-Temperature Plasma Generators," Novosibirsk, 1989*, Part II, p. 101.
12. A. V. Budin, V. A. Kolikov, B. P. Levchenko, *et al.*, Teplofiz. Vys. Temp. **32**, 628 (1994).
13. I. V. Tsvetkov, Izv. Akad. Nauk, Ser. Fiz. **58** (10), 156 (1994).
14. R. V. Mitin, *Properties and Diagnostics of Low-Temperature Plasma* (Nauka, Novosibirsk, 1977), p. 105.
15. F. G. Rutberg, A. A. Bogomaz, A. V. Budin, *et al.*, Izv. Akad. Nauk, Énerg., No. 1, 100 (1998).
16. V. P. Ignatko and G. M. Chernyavskiĭ, in *Proceedings of the I All-Union Workshop on the Dynamics of High-Current Arc Discharges in a Magnetic Field, Novosibirsk, 1990*, p. 88.
17. A. A. Bogomaz, V. S. Borodin, B. P. Levchenko, and F. G. Rutberg, Zh. Tekh. Fiz. **47**, 121 (1977) [Sov. Phys. Tech. Phys. **22**, 68 (1977)].
18. I. A. Glebov and F. G. Rutberg, *High-Power Plasma Generators* (Énergoatomizdat, Moscow, 1985).
19. V. P. Ignatko, in *Proceedings of the IV All-Union Conference on Physics of Low-Temperature Plasma, Kiev, 1975*, Part II, p. 17.
20. A. A. Bogomaz, A. V. Budin, V. A. Kolikov, *et al.*, Izv. Akad. Nauk, Énerg., No. 1, 64 (1998).
21. S. Yu. Luk'yanov, *High-Temperature Plasma and Controlled Nuclear Fusion* (Nauka, Moscow, 1975).
22. G. A. Lyubimov, *Experimental Investigations of Plasmatrons* (Nauka, Novosibirsk, 1977), pp. 207–226.
23. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Nauka, Moscow, 1966; Academic, New York, 1966).
24. S. Crist, P. M. Sherman, and D. R. Glass, AIAA J. **4**, 68 (1966).
25. B. V. Zamyshlyaev, E. L. Stupitskiĭ, A. G. Guz', and V. G. Zhukov, *Composition and Thermodynamic Functions of Plasma: Handbook* (Énergoatomizdat, Moscow, 1984).

Translated by N. Larionova

Technical Physics, Vol. 47, No. 1, 2002, pp. 34–37. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 36–39.

SOLIDS

Splay–Splay Transition in Bistable Nematic Liquid Crystals

V. I. Tsoy

Chernyshevsky State University, ul. Astrakhanskaya 83, Saratov, 410026 Russia

e-mail: TsoyVI@info.sgu.ru

Received December 7, 2000

Abstract—The bistability of a nematic liquid crystal layer with the asymmetric directions of easy orientation and weak orientational anchoring is considered within the framework of the continuum theory. The feasibility of passive storage applications based on two differing states of lateral bending is demonstrated. Switching between these splay states does not require the motion of the domain walls and can be accomplished by the two-frequency control of the low electric field applied to the crystal. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

An important use of liquid crystals is in displays with passive storage of writable images. Bistable liquid crystals attractive for this purpose may have molecular ordering of different type. These are ferroelectric smectics [1], cholesterics [2], and nematics [3, 4]. The bistability of a nematic layer may be realized by various means. For instance, the stable states may differ in the azimuth of the preferred molecule orientation [3]. These two azimuths correspond to the directions of easy orientation on each of the specially prepared substrates. Here, switching between the orientations proceeds simultaneously over the layer plane and is caused by the chiral ion flows toward the substrates. Another way is the use of topologically independent orientational structures of lateral bending (horizontal *H* configuration) and buckling (vertical *V* configuration) in a cell with the directions of easy orientation oppositely turned relative to the substrates in the vertical plane [4]. In this case, switching between these states occurs in the electric field, which triggers the motion of the domain walls separating the *H* and *V* regions. Finally, switching may involve the thermal collapse of one of the structures (states), followed by cooling to change the material to the other state. In this study, the bistability of a nematic layer with the weak orientational anchoring of molecules on the surface is considered. It is shown that, with the anchoring parameter and the easy orientation angles properly selected, the two optically distinguishable nonhorizontal splay configurations of lateral bending can be taken as equistable states. Switching between them may be caused by the electric field applied normally to the layer and does not require the motion of domain walls.

A BISTABLE NEMATIC LAYER WITH WEAK ORIENTATIONAL SURFACE ANCHORING

Let us consider a horizontal nematic liquid crystal layer ($-L/2 < z < L/2$). The directions of easy orientation are taken to be near-horizontal on the lower substrate and near-vertical on the upper substrate. Such directions can be obtained by opposite rotations relative to the substrates in the vertical plane; i.e., the pretilt angles of the director relative to the substrates have the opposite signs: $\Theta^-(-L/2) = \Theta^- < 0$, $\Theta^+(L/2) = \Theta^+ > 0$. We also assume that the anchoring potentials on both substrates are the same.

First, we will trace the director field relaxation from the undistorted nematic state N_h with the uniform horizontal orientation (see figure). This orientation is achievable in liquid crystals with negative dielectric anisotropy in a strong electric field normal to the layer or in liquid crystals with positive dielectric anisotropy in a tangential electric field. When the field is off, the directors at the layer boundaries relax to the directions of easy orientation. When the surface anchoring is strong, the well-known *H* configuration, which is characterized by the interior plane $z = z_h$ where the director is horizontal, results [4]. As the surface anchoring weakens, the torques that tend to preserve the nematic order in the bulk turn the directors from the easy orientations to the horizontal direction. On the lower substrate, where the easy orientation is almost horizontal, the director may turn away from the substrate to the horizontal position and even farther. As a result, the tilt angle of the director becomes positive across the layer thickness. This corresponds to the nonhorizontal splay configuration of the lateral bending S_h.

1063-7842/02/4701-0034$22.00

Other orientational states result from the relaxation of the undistorted nematic vertical N_v configuration. It is known [5–7] that, when the anchoring between the director and the substrates is strong, either the planar vertical V configuration or the half-turn T configuration twisted about the vertical axis appears. The V state is characterized by the presence of the interior plane $z = z_v$ where the director is normal to the layer plane. The transition between the vertical and the twisted configurations is made possible via continuous straining, while the horizontal structure is topologically incompatible with both. At a weak surface anchoring, the directors on the substrates turn from the easy orientations towards the vertical. On the upper substrate, where the easy orientation is almost vertical, the director may turn away from the substrate to the vertical position and even farther. Then, the tilt angle of the director becomes negative across the layer and the nonvertical, rather than vertical, splay configuration of the lateral bending S_v is set.

In what follows, the S_h and S_v states will be considered within the framework of the continuum theory. First, we assume that the orientational elasticity is isotropic and the surface energy of the director is described by the Rapini potential [8]. Then, the free energy per unit square can be expressed as

$$F = \int (K/2)(\partial\Theta/\partial z)^2 dz + (W/2)\sin^2(\Theta_- - \Theta^-) + (W/2)\sin^2(\Theta_+ - \Theta^+), \quad (1)$$

where K is the Oseen–Frank elasticity constant; W is the surface anchoring energy; Θ^- and Θ^+ are the tilt angles of easy orientation at the boundaries $z = -L/2$ and $z = L/2$, respectively; and Θ_- and Θ_+ are the actual tilt angles of the director at the boundaries.

The variational conditions for the minimum of energy (1) are

$$\partial\Theta/\partial z = (\Theta_+ - \Theta_-)/L, \quad (2)$$

$$-K(\partial\Theta_-/\partial z) + W\sin(\Theta_- - \Theta^-)\cos(\Theta_- - \Theta^-) = 0, \quad (3)$$

$$K(\partial\Theta_+/\partial z) + W\sin(\Theta_+ - \Theta^+)\cos(\Theta_+ - \Theta^+) = 0. \quad (4)$$

From Eq. (2), it follows that the strain of the nematic order is uniform. Therefore, the two force moments acting on the directors from the substrates are equal in magnitude:

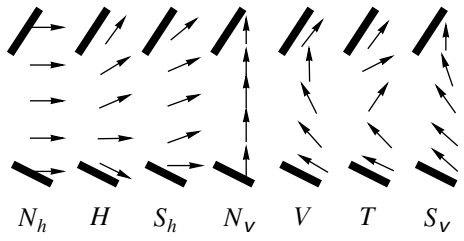
$$W\sin(2\chi^+) = -W\sin(2\chi^-), \quad (5)$$

where $\chi^- = \Theta_- - \Theta^-$ and $\chi^+ = \Theta_+ - \Theta^+$ are the tilt angles of the director with respect to the directions of easy orientation.

Rewriting Eqs. (3) and (4) in terms of (5), we obtain

$$(2K/WL)(\chi^+ - \chi^- + \Theta^+ - \Theta^-) + \sin 2\chi^+ = 0, \quad (6)$$

$$(2K/WL)(\chi^+ - \chi^- + \Theta^+ - \Theta^-) - \sin 2\chi^- = 0. \quad (7)$$



Possible configurations in a bistable nematic layer.

Equations (6) and (7) admit two solutions that differ in sign of the director tilts on the substrates. The tilt angles relate as $\chi^- = -\chi^+ > 0$ and $\chi^- = -\chi^+ < 0$ for one and the other solution, respectively. Continuous strain (2) is possible if $\Theta_+ - \Theta_- < \pi$. Therefore, at $\Theta^+ - \Theta^- = \pi/2$, the second solution can readily be obtained from the first one by changing the sign of the angles between the director and the directions of easy orientation on both substrates and reversing the direction of easy orientation on one of the substrates. For example, with the anchoring parameter $K/WL = 0.1$ and the angles of easy orientation $\Theta^- = -5°$ and $\Theta^+ = 85°$, the respective tilt angles of the director on the substrates are $\Theta_- = 2.6°$, $\Theta_+ = 77.4°$ and $\Theta_- = -12.6°$, $\Theta_+ = -87.4°$ for the two possible configurations. The former is the S_h state, and the latter is the S_v state. These configurations are feasible for passive storage applications. Indeed, in the first configuration, the tilt angle of the director spans the range $(\Theta_+ - \Theta_-) = 2\chi^+ + \pi/2$, and, in the second one, the range is the same but with the opposite sign. This means that strains (2), as well as the director tilts to the directions of easy orientation on the substrates, differ only in sign. Consequently, energy (1) is the same for both equilibrium states and they are equally stable.

SPLAY–SPLAY TRANSITION DYNAMICS

We will now consider the conventional dynamic model for the S_h and S_v states of the bistable layer, as well as switching between them in an external electric field, within the context of the continuum theory. It is assumed that the director is inertialess, the rotational viscosity is isotropic, the fluid flow is absent [9], and the anisotropic energy of surface anchoring is described by a function close to the Rapini potential [10, 11]. We take into account the anisotropic orientational viscosity and the nonuniformity of the field applied to the strained nematic layer [12]. The surface viscosity is also taken into consideration [13, 14].

Within the model, the director tilt is defined by the dynamic equilibrium of the torques:

$$\gamma_1(\partial\Theta/\partial t) = (\partial/\partial z)(\partial F_v/\partial\Theta_z) - \partial F_v/\partial\Theta, \quad (8)$$

$$\gamma_1(\partial\varphi/\partial t)\cos^2\theta = (\partial/\partial z)(\partial F_v/\partial\varphi_z) - \partial F_v/\partial\varphi, \quad (9)$$

$$\gamma_s(\partial\Theta_-/\partial t) = \partial F_v/\partial\Theta_z - \partial F_s/\partial\Theta_-, \quad (10)$$

$$\gamma_s(\partial\Theta_+/\partial t) = -\partial F_v/\partial\Theta_z - \partial F_s/\partial\Theta_+. \quad (11)$$

Here, γ_1 and γ_s are the volume and surface rotational viscosities, respectively; F_v and F_s are the volume and surface energy densities, respectively; and Θ_z and φ_z are the derivatives normal to the layer. The volume energy density is given by

$$F_v = F_d + F_e, \quad (12)$$

where

$$F_d = (K_{11}\cos^2\Theta + K_{33}\sin^2\Theta)\Theta_z^2/2 + \cos^2\Theta(K_{22}\cos^2\Theta + K_{33}\sin^2\Theta)\varphi_z^2/2, \quad (13)$$

$$F_e = -\varepsilon_0\varepsilon_2 U^2\beta^2/2\left(\int\beta dz\right)^2. \quad (14)$$

Here, K_{11}, K_{22}, and K_{33} are the orientational viscosity constants; $\beta = 1/(1 + \delta\varepsilon\sin^2\Theta)$; $\delta\varepsilon = (\varepsilon_1 - \varepsilon_2)/\varepsilon_2$; ε_1 and ε_2 are the permittivities parallel and perpendicular to the director, respectively; and U is the root-mean-square voltage across the substrate electrodes.

The anisotropic potential of the director anchoring energy is close to the Rapini potential [10] and is given by

$$F_s = (W/2)\sin^2(\Theta_- - \Theta^-) - (W_4/4)\sin^4(\Theta_- - \Theta^-) + (W/2)\sin^2(\Theta_+ - \Theta^+) - (W_4/4)\sin^4(\Theta_+ - \Theta^+). \quad (15)$$

The numerical solution of Eqs. (8)–(11) obtained with the typical material constants confirms the supposition on the S_h and S_v stable states in the layer. It turns out that an electric field applied transversely to the layer may cause switching between these states that takes place simultaneously over the plane of the layer.

We used the material constants close to those in [14–16]: $K_{11} = 1 \times 10^{-11}$ N, $K_{22} = 0.5 \times 10^{-11}$ N, $K_{33} = 2 \times 10^{-11}$ N, $\gamma_1 = 0.1$ N cm^{-2}, $\gamma_s = 3 \times 10^{-8}$ N cm^{-1}, $\varepsilon_1 = 15$, and $\varepsilon_2 = 5$. For the 3-μm-thick layer, the surface viscosity parameter $\gamma_1 L/\pi\gamma_s$ was ≈3. The surface anchoring parameter $\pi K_{11}/WL$ was assumed to be equal to unity, which gives for the anchoring energy $W \sim 1 \times 10^{-5}$ J m^{-2}. The correction to the Rapini potential was taken to be equal to $W_4/W = 0.5$ [10]. The numerical integration of Eqs. (8)–(11) (for the easy orientation angles $\Theta^- = -15°$ and $\Theta^+ = 78°$) yielded two distinct configurations of lateral bending with the same energies per unit surface area: $F_v \sim 3 \times 10^{-6}$ J m^{-2}. The respective tilt angles of the director on the substrates are $\Theta_- = 3°$, $\Theta_+ = 50°$ and $\Theta_- = -35°$, $\Theta_+ = -73°$. These states differ in optical thickness. With the refractive indices $n_1 = 1.65$ and $n_2 = 1.5$, the difference amounts up to 0.2 μm. According to our calculations, switching between the S_h and S_v states can be induced by a transverse electric field at frequencies corresponding to the different signs of the dielectric anisotropy. The electric voltage was 3 V, and the transition from positive to negative dielectric anisotropy was accomplished by interchanging the permittivities. S_h-to-S_v switching requires the application of the field for which dielectric anisotropy is positive. In this case, the director tilt to the layer plane grows and, at some time instant, reaches the critical value $\Theta_- = \Theta^- + \pi/2$ on the lower substrate. At this value, the director is perpendicular to the direction of easy orientation. Subsequently, anchoring on the lower substrate assists, rather than prevents, the field-induced rotation of the director up to the vertical position. This gives rise to the vertical *V* state, which relaxes, through the continuous variation of the director configuration, to the stable S_v state when the field is off. For reverse S_v-to-S_h state switching, the external field frequency should be varied so that the dielectric anisotropy becomes negative. As a result, the director tilt to the layer plane decreases. At a sufficiently high voltage, the director at the upper substrate deviates from the direction of easy orientation by the critical angle $\Theta_- = \Theta^- + \pi/2$, becoming perpendicular to this direction. Then, anchoring on the upper substrate assists, rather than prevents, the field-induced rotation of the director up to the horizontal position. This gives rise to the horizontal *H* configuration, which relaxes to the stable S_h state when the field is off. Typical time intervals required for the switching were estimated at 300 ms.

CONCLUSION

The configuration of the director field in a bistable liquid layer strongly depends on the orientational anchoring at its surfaces. Varying the parameters and the geometry of the cell, one may choose two equilibrium states that have the same elastic energy and, thus, meet the requirements for passive storage applications. It is significant that, at a finite anchoring energy, switching in the bistable cell may be induced by a low uniform (over the layer plane) electric field without the motion of domain walls. Our results suggest that the equilibrium states may be two splay configurations that arise from the horizontal and vertical states after their relaxation. Switching between the *H* and *V* states, as well as between the *H* and S_v states, proceeds by a similar scenario.

REFERENCES

1. N. A. Clark and S. T. Lagerwall, Appl. Phys. Lett. **36**, 899 (1980).
2. D. W. Berreman and W. R. Heffner, Appl. Phys. Lett. **37**, 109 (1980).
3. R. Barberi and G. Durand, Appl. Phys. Lett. **58**, 2907 (1991).
4. G. D. Boyd, J. Cheng, and P. D. T. Ngo, Appl. Phys. Lett. **36**, 556 (1980).

5. G. Porte and J. P. Jadot, J. Phys. (Paris) **39**, 213 (1978).
6. J. Cheng, R. N. Thurston, and D. W. Berreman, J. Appl. Phys. **52**, 2756 (1981).
7. L. Komitov, G. Hauch, and H. D. Koswig, Phys. Status Solidi A **97**, 645 (1986).
8. J. Nehring, A. R. Kmetz, and T. J. Sheffer, J. Appl. Phys. **47**, 850 (1976).
9. P. Pieransky, F. Brochard, and E. Gyuon, J. Phys. (Paris) **34**, 35 (1973).
10. H. Yokoyama and H. A. van Sprang, J. Appl. Phys. **57**, 4520 (1985).
11. L. M. Blinov, D. Z. Radzhaboy, D. B. Subachyus, and S. V. Yablonskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **53**, 223 (1991) [JETP Lett**. 53**, 238 (1991)].
12. H. J. Deuling, Mol. Cryst. Liq. Cryst. **27**, 81 (1974).
13. V. I. Tsoy, Mol. Cryst. Liq. Cryst. **264**, 51 (1995).
14. G. Derfel and B. Gajewska, Liq. Cryst. **22**, 297 (1997).
15. H. A. van Sprang and H. G. Koopman, J. Appl. Phys. **64**, 4873 (1988).
16. S. Yablonski, M. Rajteri, S. Oldano, and G. Durand, Proc. SPIE **2731**, 87 (1996).

Translated by A. Sidorova-Biryukova

Technical Physics, Vol. 47, No. 1, 2002, pp. 38–48. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 40–50.

SOLIDS

Magnetoelastic Waves in an In-Plane Magnetized Ferromagnetic Plate

Yu. A. Filimonov and Yu. V. Khivintsev

Institute of Radio Engineering and Electronics, Saratov Branch, Russian Academy of Sciences, Saratov, 410019 Russia

e-mail: khiv@sfire.san.ru

Received March 23, 2001

Abstract—The spectrum of magnetoelastic waves propagating along the magnetic field in an in-plane magnetized ferromagnetic plate is numerically investigated in the exchangeless approximation. No restrictions are imposed either on the field pattern of backward volume magnetostatic waves (BVMSWs) or elastic waves supported by a plate of a given geometry across the plate or on the relationship between the sound velocity v_S and the phase velocity of the magnetoelastic waves $v = \omega/q$ (ω is the frequency, q is the wave number). The resonance interaction of the BVMSWs and elastic waves is accompanied, as a rule, by the formation of "stop" bands $\delta\omega$ that are proportional to the magnetoelastic coupling constant b. When the BVMSWs are in resonance with Lamb and shear elastic modes the values of the magnetoelastic gaps $\delta\omega$ at $v \approx v_S$ turn out to be of the same order. For $v \gg v_S$, the efficiency of the interaction between the BVMSWs and transverse Lamb modes is almost one order of magnitude higher. If the frequency spacing $\Delta\omega$ between the elastic modes is smaller than the magnetoelastic gap in the spectrum ($\Delta\omega \le \delta\omega$), which takes place, particularly, in the region of crowding the elastic mode spectrum ($v \approx v_S$), the resonant interaction results in mixing the dispersion laws for the elastic modes. Namely, a surface mode may transform into a volume one and a shear mode, into the Lamb mode or into a shear mode with another number. The resonance interaction of the shear and Lamb elastic modes not only forms the magnetoelastic gaps $\delta\omega \sim b^2$ but also changes the efficiency of elastic wave coupling with the magnetic subsystem. This may show up as the coexistence of the effects of "repulsing" both the dispersion laws and the damping decrements of the elastic waves at the resonance frequency. It is shown that magnetostriction splits the cutoff frequencies of both transverse Lamb modes and shear modes, as well as the long-wave ($q \longrightarrow 0$) frequency limits f_0 of the BVMSW modes. This may cause the resonance interaction between BVMSW modes of equal evenness in a narrow frequency band $\Delta \sim b$ near f_0. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The finding of the magnetoelastic wave (MEW) spectrum is of critical importance in studying the excitation and propagation of both spin and elastic waves in ferromagnetics [1]. In ferromagnetic layers and thin-film structures, the MEW spectrum has been studied, as a rule, for the ranges of frequency ω and wave numbers q that correspond to the propagation either of surface acoustic waves (SAWs) or uniform volume sound (when the MEW phase velocity $v = \omega/q$ is of the order of the sound velocity in the medium v_S) or of magnetostatic waves (MSWs) ($v \gg v_S$) [1]. However, the formation of the MEW spectrum in ferromagnetic layers without any restrictions imposed on the relationship between v and v_S should be known for the construction of the full MEW spectrum. The aim of this work is to study the MEW spectrum in an in-plane magnetized ferromagnetic plate where the waves propagate along the magnetic field $\mathbf{H}_0$.

It should be noted that the MEW spectrum in a ferromagnetic plate for a given magnetizing field configuration has been studied in [2]. In that work, the resonance interaction of BVMSWs with the elastic modes of the plate, which are nearly uniformly distributed across the plate if $v \approx v_S$, was examined. In our study, the MEW spectrum of an in-plane magnetized ferrite plate is analyzed with a numerical approach, so that no restrictions are imposed on the distribution of the wave fields across the plate. Therefore, we can consider the interaction of BVMSWs not only with SAWs and volume modes of lower numbers but also with volume elastic modes near their cutoff frequencies. In the latter case, so-called "fast" MEWs with $v \gg v_S$ form [1]. Moreover, the use of the numerical approach allowed us to abandon the commonly used analytical approximation of "weak" magnetoelastic coupling of waves, which implies that the frequency difference $\Delta\omega$ between the elastic and magnetostatic modes in the spectrum of the plate far exceeds the "spacing" between the dispersion curves $\delta\omega$ at the phase synchronism frequency, $\Delta\omega \gg \delta\omega$. The resonance interaction of the magnetostatic and elastic waves at strong magnetoelastic coupling is considered for the case $v \approx v_S$, where the volume elastic modes crowd together and the condition $\Delta\omega \le \delta\omega$ holds for the magnetic and elastic parameters typical of ferrite plates. Finally, we consider magnetostriction-related effects in the wave spectra

that appear far away from the points of intersection between the dispersion curves of the BVMSWs and elastic waves. These effects are the resonance interaction of Lamb and shear elastic modes, as well as the splitting of the cutoff frequencies (long-wave frequency limits) for dipole BVMSWs and purely shear modes and for the transverse Lamb modes.

STATEMENT OF THE PROBLEM AND THE DISPERSION RELATION

We will consider waves propagating in an isotropic ferromagnetic plate lying in the $x0z$ plane and occupying the space $-d/2 \le y \le d/2$. Let them propagate along the $0x$ axis aligned with an external magnetic field $\mathbf{H}_0$ (insert in Fig. 1). It is assumed that the magnetostatic and exchangeless approximations are valid [1].

In this case, the magnetostatics equations and the linearized equations of motion of magnetization and elasticity in a ferrite plate can be represented in the form [2]

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + 4\pi\frac{\partial m_y}{\partial y} = 0,$$

$$\frac{\partial m_y}{\partial t} = -\gamma H_0 m_z - \gamma M_0^2 b\frac{\partial u_z}{\partial x},$$

$$\frac{\partial m_z}{\partial t} = \gamma H_0 m_y - \gamma M_0\frac{\partial\psi}{\partial y} + \gamma M_0^2 b\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right),$$

$$\frac{\partial^2 u_x}{\partial t^2} = \frac{M_0 b}{\rho}\frac{\partial m_y}{\partial y} + v_l^2\frac{\partial^2 u_x}{\partial x^2} + v_t^2\frac{\partial^2 u_x}{\partial y^2} + (v_l^2 - v_t^2)\frac{\partial^2 u_y}{\partial x\partial y}, \quad (1)$$

$$\frac{\partial^2 u_y}{\partial t^2} = \frac{M_0 b}{\rho}\frac{\partial m_y}{\partial x} + v_t^2\frac{\partial^2 u_y}{\partial x^2} + v_l^2\frac{\partial^2 u_y}{\partial y^2} + (v_l^2 - v_t^2)\frac{\partial^2 u_x}{\partial x\partial y},$$

$$\frac{\partial^2 u_z}{\partial t^2} = \frac{M_0 b}{\rho}\frac{\partial m_z}{\partial x} + v_t^2\left(\frac{\partial^2 u_z}{\partial x^2} + \frac{\partial^2 u_z}{\partial y^2}\right).$$

Here, ρ is the density of the material; M_0 is the saturation magnetization; v_t and v_l are the transverse and longitudinal sound velocities, respectively; γ is the gyromagnetic ratio; b is the dimensionless magnetostriction constant; ψ is the magnetostatic potential; m_y and m_z are the high-frequency components of the magnetization vector ($m_y, m_z \ll M_0$), respectively; and u_x, u_y, and u_z are the components of the displacement vector.

The solution of system (1) was taken in the form $\exp(py)\exp(i[qx - \omega t])$, where p characterizes the field distribution across the structure. The compatibility condition for system (1) results in the fourth-order equation

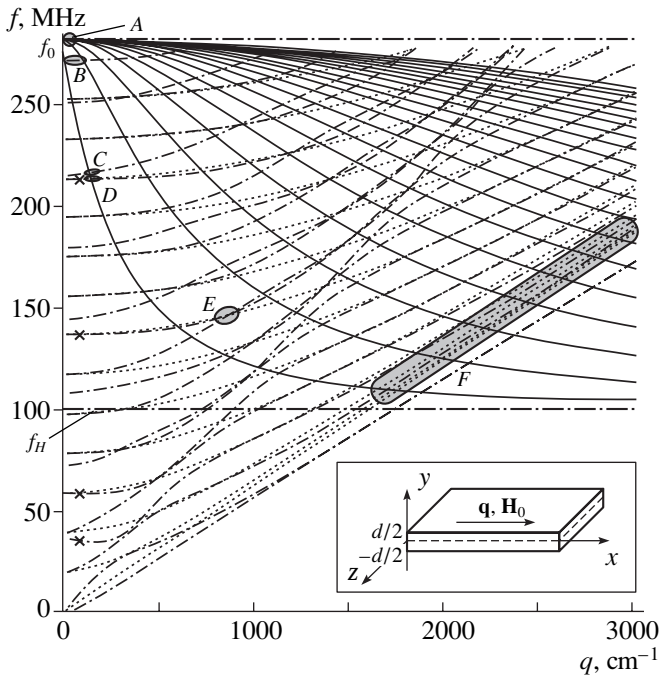


Fig. 1. Dispersion of BVMSWs (continuous curves) and shear (dashed curves) and Lamb (dash-and-dot curves) elastic modes in the absence of magnetostriction and dissipation. The geometry of the problem is shown in the insert.

in p^2, which can we written in terms of the dimensionless variables and parameters:

$$\Delta_1\Delta^2\Delta_3 + \xi Q^2(\Delta_2 A(\Delta_3(\chi^2+1) - 4\Delta_2\chi^2) + \Delta_2\Delta_3(A(\chi^2-1)+\chi^2) + \xi Q^2(\Delta_3(\chi^2+1) - 4\Delta_2\chi^2)) = 0, \quad (2)$$

where

$$\Delta_1 = (F^2 - A^2)(\chi^2 - 1) - A\chi^2,$$

$$\Delta_2 = F^2 + BQ^2(\chi^2 - 1),$$

$$\Delta_3 = F^2 + CQ^2(\chi^2 - 1),$$

$$\xi = ((4\pi)^3\rho)^{-1}(b/(\gamma d))^2,$$

$$A = H_0/(4\pi M_0), \quad B = (v_t/(\gamma 4\pi M_0 d))^2,$$

$$C = (v_l/(\gamma 4\pi M_0 d))^2, \quad F = \omega/(\gamma 4\pi M_0),$$

$$Q = qd, \quad \chi = p/q.$$

The general solution of system (1) was taken in the form

$$\psi, m_y, m_z, u_x, u_y, u_z \sim \sum_{j=1}^{8}\alpha_{jk}A_j\exp(p_j y) \times \exp(i[qx - \omega t]), \quad (3)$$

where α_{jk} ($k = \psi, m_y, m_z, u_x, u_y, u_z$) are the amplitude coefficients of the wave components, which were found from (1) by eliminating one of the equations and equating the amplitude coefficients for one of the wave components to unity.

On the surfaces of the plate ($y = \pm d/2$), solution (3) should obey the conditions of the absence of external forces [3]

$$\sigma_{zy} = \frac{\partial u_z}{\partial y} = 0, \tag{4}$$

$$\begin{aligned}\sigma_{xy} &= \frac{M_0 b}{\rho} m_y + v_t^2\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right) = 0,\\ \sigma_{yy} &= v_l^2\frac{\partial u_y}{\partial y} + (v_l^2 - 2v_t^2)\frac{\partial u_x}{\partial x} = 0\end{aligned} \tag{5}$$

and the continuity conditions for the normal components of the magnetic induction and for the tangential components of the magnetic field. The latter conditions can be written in the form [3]

$$\frac{\partial \psi}{\partial y} \mp i\frac{\partial \psi}{\partial x} + 4\pi m_y = 0, \tag{6}$$

where the upper and lower signs correspond to the upper and bottom boundaries of the plate.

Substituting (3) into (4), we obtain the system of algebraic equations in A_j. The compatibility condition for this system, which defines the relationship between ω and q at given p_j, is expressed as

$$\det(a_{ij}) = 0, \tag{7}$$

where $\det(a_{ij})$ is the eighth-order determinant with the elements

$$a_{1j} = \left(\Delta_2(\chi_j) + \frac{\xi}{A}Q^2\right)\Delta_2(\chi_j)\Delta_3(\chi_j),$$

$$a_{3j} = (\chi_j^2 - \chi_j)\left(\Delta_2(\chi_j) + \frac{\xi}{A}Q^2\right),$$

$$\begin{aligned}a_{5j} = (1-\chi_j)(F^2\Delta_3(\chi_j) - 4B(C-B)Q^4\chi_j^2)\\ \times\left(\Delta_2(\chi_j) + \frac{\xi}{A}Q^2\right),\end{aligned}$$

$$a_{7j} = (\chi_j^2 - \chi_j)\Delta_2(\chi_j)\Delta_3(\chi_j),$$

$$a_{kj} = a_{(k-1)j}\exp(\chi_j Q) \quad (k = 2, 4, 6, 8).$$

Thus, the problem of calculating the MEW spectrum is reduced to finding ω and q that simultaneously satisfy Eqs. (2) and (7). We solved it numerically following the procedure adopted in [3]. The dissipation was introduced by the standard substitution [1] $H_0 \Rightarrow H_0 - i\Delta H$, $v_l^2 \Rightarrow v_l^2(1 - i\omega\eta_{11}/c_{11})$, and $v_t^2 \Rightarrow v_t^2(1 - i\omega\eta_{44}/c_{44})$, where ΔH is the half-width of the ferromagnetic resonance (FMR) band and η_{11}, η_{44} and c_{11}, c_{44} are the viscosities and the elastic moduli of the longitudinal and transverse sounds, respectively. The damping was taken into account according to whether the problem is solved for traveling or standing waves [1]. In the first case, the real ω was fixed and complex $q = q' + iq''$ were found, where the real part, q', specifies the dispersion and the imaginary part, q'', the spatial damping decrement of the waves. In the second case, the real q was fixed by the boundary conditions and complex $\omega = \omega' - i\omega''$ were found, where ω' defines the eigenfrequencies and ω'', the damping of eigenmodes.

Determinant (5) can be represented in the form

$$\begin{vmatrix} \|S\| & \sim b & \sim b & \sim b \\ \sim b & \| & T & \sim q & \| & \sim b \\ \sim b & \| & \sim q & L & \| & \sim b \\ \sim b & \sim b & \sim b & \|M\| \end{vmatrix} = 0, \tag{8}$$

where the diagonal elements of the determinant, which are combined into the 2×2 matrices $\|S\|$, $\|T\|$, $\|L\|$, and $\|M\|$, obey the dispersion relations for purely shear ($\Delta S = 0$), Lamb ($\Delta L = TL + \delta = 0$, where $\delta \sim q^2$), and backward volume magnetostatic waves ($\Delta M = 0$), which follow from boundary conditions (4)–(6), respectively, in the absence of magnetoelastic coupling. The elements $\sim b$ of the determinant are responsible for magnetoelastic interaction in the plate and make these waves coupled. Determinant (8) can be represented in the form

$$\Delta S\Delta L\Delta M = \delta_1\Delta S + \delta_2\Delta L + \delta_3\Delta M + \delta_4, \tag{9}$$

where δ_1, δ_2, δ_3, δ_4 are responsible for wave interaction and contain the terms $\sim I^{2n}$, where I is the overlap integral for the interacting wave fields, and n ($1 \le n \le 4$) characterizes the magnetoelastic coupling order.

The overlap integral, in turn, can be expressed through the magnetization $\mathbf{m}(y)$ of one wave and the magnetoelastic field $\mathbf{h}_{me}(y)$ of the other:

$$I = \int_{-d/2}^{d/2} \mathbf{m}(y)\mathbf{h}_{me}(y)dy. \tag{10}$$

It is clear from (10) that when the magnetostatic and elastic waves are in resonance, $I \sim b$, and the coupling terms δ_1, δ_2, responsible for the interaction of the Lamb and shear elastic waves with the BVMSWs, have the order of b^2. The resonance of the Lamb and shear elastic waves is characterized by the coupling term $\delta_3 \sim b^4$ ($I \sim b^2$).

In our case, the symmetry about the plane $y = 0$ sets coupling only between the modes that have the equal evenness of, respectively, $\mathbf{m}(y)$ and $\mathbf{h}_{me}(y)$. Taking into account the distribution of the high-frequency magnetization $\mathbf{m}(y)$ in the BVMSW modes [4] and of the displacements $\mathbf{u}(y)$ in the elastic modes [5], as well as con-

sidering their coupling with the magnetoelastic fields $\mathbf{h}_{me}$ in Eqs. (1), one can show that the BVMSW modes interact with the shear and longitudinal Lamb elastic modes that have the same evenness, as well as with the transverse Lamb modes of opposite evenness.

RESULTS OF CALCULATIONS

The calculations were performed for a plate of thickness $d = 100$ μm and parameters typical of Ga, Sc-substituted yttrium iron garnet (Ga, Sc : YIG): $4\pi M_0 = 250$ G, $\Delta H = 0.2$ Oe, $\gamma/(2\pi) = 2.8$ MHz/Oe, $\rho = 5.17$ g/cm^3, $v_l = 7.11 \times 10^5$ cm/s, $v_t = 3.85 \times 10^5$ cm/s, $\eta_{11}/c_{11} = 2.4 \times 10^{-15}$ s, $\eta_{44}/c_{44} = 3.0 \times 10^{-16}$ s, and $b = 1600$. The external magnetic field (unless otherwise specified) was taken to be equal to $H_0 = 35$ Oe.

Before turning to the discussion of the results obtained, we note that in the absence of magnetostriction ($b = 0$), BVMSW modes [4] and elastic modes [5] can propagate in such a plate. Figure 1 shows the dispersion of these waves in the absence of dissipation. The continuous curves refer to the BVMSW modes, which occupy the frequency ($f = \omega/2\pi$) range $f_H \le f \le f_0$, where $f_H = \gamma H_0/(2\pi)$ and $f_0 = \sqrt{f_H^2 + f_H f_M}$ are the short- and long-wave frequency limits of the BVMSW spectrum and $f_M = \gamma 4\pi M_0/(2\pi)$. In the dipole approximation, the cutoff frequencies of the BVMSW modes with a different number m of half-waves falling into the plate thickness are coincident with the frequency f_0.

The Lamb modes with displacements polarized in the sagittal plane $x0y$ are shown by the dash-and-dot lines. The dispersion of the Lamb modes is positive, as a rule. However, some of the Lamb modes have the domains of negative dispersion at small wave numbers q (crosses in Fig. 1). The cutoff frequencies of the Lamb modes are given by

$$f_L = \frac{v_l l}{2d}, \quad f_T = \frac{v_t t}{2d}, \qquad (11), (12)$$

where l and t are the numbers of the longitudinal and transverse Lamb modes, respectively, that coincide with the number of half-waves falling into the thickness of the plate.

The Lamb modes whose dispersion curves originate at the cutoff frequencies defined by (11) are called longitudinal modes. Those whose dispersion curves originate at frequencies (12) are called transverse modes.[1]

The dispersion curves of the purely shear elastic modes in the plate with displacements polarized along the 0z axis are depicted by the dashed lines. The dispersion of the shear elastic modes is positive, and the cutoff frequencies f_S characterized by the number s of half-waves falling into the plate thickness turn out to be coincident with the cutoff frequencies of the transverse Lamb modes having the number $t = s$ and are obtained from (12).

Magnetostriction ($b \neq 0$) results in the interaction of the BVMSW with the elastic waves, producing the general spectrum of the hybrid MEWs. Accordingly, the dispersion and the decay of the interacting waves change, the changes occurring not only at the frequencies of wave synchronism but also far away from the points of intersection between the dispersion curves of the BVMSWs and the elastic waves. Let us consider the most prominent changes in the spectrum.

SPLITTING OF THE BVMSW AND ELASTIC MODE CUTOFF FREQUENCIES

With magnetoelastic coupling, the cutoff frequencies of the shear and transverse Lamb elastic modes with the equal numbers ($s = t$), as well as of BVMSW modes of various numbers m, split at the frequency f_0; i.e., degeneration is removed (areas A and B, respectively, in Fig. 1). Splitting, in this case, is achieved through the interaction of the transverse Lamb modes with the BVMSW modes and appears as the offsets of the cutoff frequencies from the initial position by Δ_e and Δ_m (Fig. 2). These offsets can be calculated using the expression for the spectrum of MEWs propagating normally to the magnetizing field in an infinite ferromagnetic provided that the demagnetizing field and the restriction imposed on the wave number $q = \pi n/d$ [6] ($n = s = t = m$) by the boundary conditions are taken into account:

$$\Delta_m = \sqrt{\frac{1}{2}(f_0^2 + f_T^2 \pm R)} - f_0,$$

$$\Delta_e = \sqrt{\frac{1}{2}(f_0^2 + f_T^2 \mp R)} - f_T,$$

where

$$R = \sqrt{(f_0^2 - f_T^2)^2 + (M_0 b n/d)^2 f_H f_M/(4\pi\rho)},$$

f_T is determined from (12), and the upper and lower signs refer to the modes with the numbers $n < 2df_0/v_t$ and $n > 2df_0/v_t$, respectively.

The dependences $\Delta_e(n)$ and $\Delta_m(n)$ calculated without regard for dissipation in the system are shown in the insert to Fig. 2a. The offsets are maximal for the modes with the numbers $n = 14$ and 15, for which the cutoff frequencies f_T lie most closely to the frequency f_0 and, hence, the elastic waves are most closely approaching the phase synchronism with the BVMSW modes. At $n \longrightarrow \infty$, the elastic mode offset $\Delta_e \longrightarrow 0$, whereas the BVMSW mode remains finite and, in our case, reaches

[1] Far away from the cutoff frequencies, the Lamb modes have the transverse and longitudinal components simultaneously and the relationship between them may vary. To avoid confusion, we hereafter will call them, as before, transverse and longitudinal according to which cutoff frequencies are the origin of their dispersion curves.

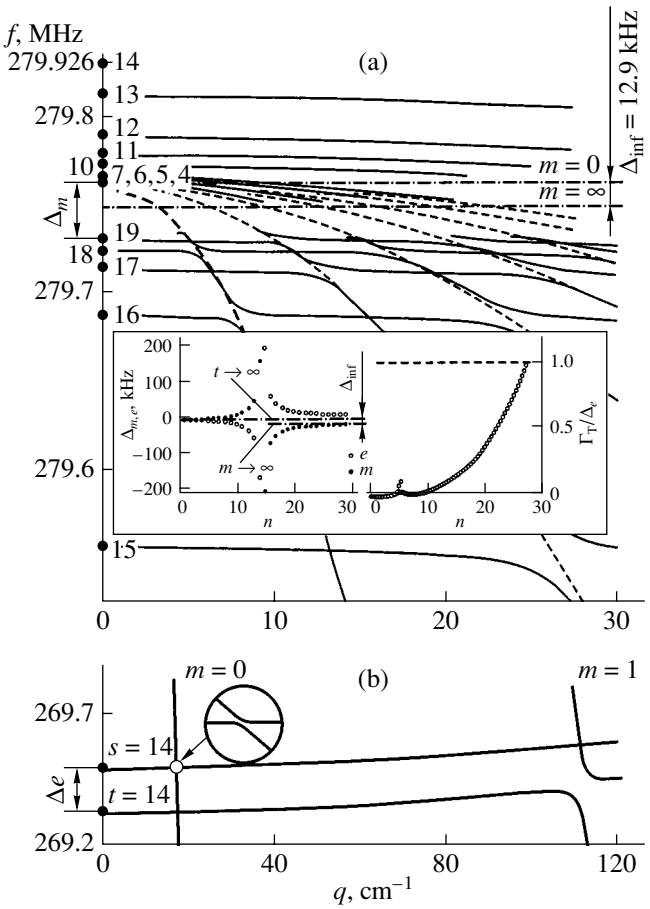


Fig. 2. (a) Dispersion of BVMSWs and (b) transverse elastic modes with the numbers $s = t = 14$ near their cutoff frequencies with magnetostriction and without dissipation. Dashed curves, BVMSW dispersion without magnetostriction. The cutoff frequencies are indicated by circles with the mode numbers alongside. The insert shows the offsets of the frequencies of the BVMSWs (Δ_m) and transverse Lamb modes (Δ_e) vs. mode number, as well as the ratio of the frequency decrement to the offset for the Lamb mode (Γ_T/Δ_e).

the value $\Delta_m(n \longrightarrow \infty) = \Delta_{inf} \approx 13$ kHz (Fig. 2a). Note also that, as the cutoff frequencies of the BVMSW modes split, these modes begin to resonantly interact. The interaction obeys the selection rule imposed by the symmetry of the problem (see the behavior of the continuous and dashed curves in Fig. 2a).

The splitting of the cutoff frequencies of the BVMSW modes and the elastic modes is observed if

$$\Delta'_m(n) - \Delta'_m(n+1) > \max\{\Gamma_M(n), \Gamma_M(n+1)\}, \quad (13)$$

$$\Delta'_e(n) > \max\{\Gamma_S(n), \Gamma_T(n)\}, \quad (14)$$

where $\Delta'_m(n)$ and $\Delta'_e(n)$ are the offsets of the BVMSW modes and Lamb modes, respectively, with consideration for losses and $\Gamma_M(n)$, $\Gamma_S(n)$, and $\Gamma_T(n)$ are the imaginary parts of the frequency $(\omega''/(2\pi))$ of the BVMSW modes, shear modes, and Lamb modes, respectively. Since the losses are small ($\omega'' \ll \omega'$), we can write

$$\Gamma_S = \pi\eta f_S^2,$$

$$\Gamma_M = \Gamma_M^0 - \frac{\Gamma_M^0 - \Gamma_T^0}{2}(1-t),$$

$$\Gamma_T = \Gamma_M^0 + \frac{\Gamma_M^0 - \Gamma_T^0}{2}(1-t),$$

where $\Gamma_M^0 = (\gamma\Delta H/(2\pi))(f_H + f_M/2)/f$, $\Gamma_T^0 = \Gamma_S$, and $t = |f_0^2 - f_T^2|/R$.

For the given dissipation parameters, the values $\Delta'_m(n)$ and $\Delta'_e(n)$ differ from those calculated without regard for the losses by less than 2%. The damping of the BVMSW modes $\Gamma_M(n)$ depends largely on intrinsic relaxation processes. For the elastic modes, along with intrinsic viscous losses, magnetic losses due to magnetostriction are also appreciable (for modes with numbers $n < 20$, the magnetic losses dominate over the viscous ones for the given parameters). Accordingly, for condition (14) to be met, the Lamb elastic modes must decay. Note that if condition (13) does not hold for any of the BVMSW modes, (14) holds for elastic modes with numbers $n < 80$ (insert in Fig. 2).

RESONANCE INTERACTION OF BVMSWs AND ELASTIC MODES

Under the conditions of phase synchronism, to which the points of intersection of the dispersion curves in Fig. 1 correspond, BVMSWs and the elastic waves interact resonantly. With the former, the symmetric and antisymmetric SAWs, as well as the purely shear and transverse Lamb modes with numbers $s, t \le 15$ and also the longitudinal Lamb modes with numbers $l \le 7$, can interact in a resonant fashion (Fig. 1). We will consider the interaction of the elastic mode only with two first BVMSW modes ($m = 0, 1$).

The resonance interaction of the BVMSWs with the elastic modes for which the cutoff frequencies lie below the short-wave frequency limit of the spectrum f_H occurs at the phase velocities $v \approx v_S$ (area F in Fig. 1). The resonances of the BVMSWs with elastic modes whose cutoff frequencies fall into the frequency band $f_H \le f \le f_0$ are observed in the region of the "fast" magnetoelastic waves, where $v \gg v_S$.

Note that if magnetoelastic coupling is weak, the set of resonances disintegrates into the resonances of the BVMSW with the purely shear modes or with the Lamb modes. Then, Eq. (9) is reduced to the pair of equations of the form

$$\Delta L \Delta M = \delta_1 \quad (15)$$

for the resonance of the BVMSW with the Lamb modes or

$$\Delta S \Delta M = \delta_2 \tag{16}$$

for the resonance of the BVMSW with the purely shear modes.

Assuming that the deviations of the frequency, $\Delta\omega$, and of the wave number, Δq, from the resonance points (ω^*, q^*) are small ($\Delta\omega \ll \omega$, $\Delta q \ll q$) and using the standard procedure [1], we can reduce Eqs. (9) to the form

$$\begin{aligned}&(\Delta\omega - \Delta q V^g_{\mathrm{M}} + iQ_{\mathrm{M}}V^g_{\mathrm{M}})\\ &\times(\Delta\omega - \Delta q V^g_{\mathrm{L(T,S)}} + iQ_{\mathrm{L(T,S)}}V^g_{\mathrm{L(T,S)}}) = \delta\omega^2_{\mathrm{me}},\end{aligned} \tag{17}$$

where V^g_{M} and $V^g_{\mathrm{L(T,S)}}$ are the group velocities of the BVMSW and the elastic wave, respectively, at the resonance frequency; Q_{M} and $Q_{\mathrm{L(T,S)}}$ are the spatial damping decrements of the BVMSW and the elastic wave, respectively, at the resonance frequency (with regard for the coupling between the magnetic and elastic subsystems); and $\delta\omega^2_{\mathrm{me}} = \delta_{1(2)}(\partial\Delta M/\partial\omega\partial\Delta L(S)/\partial\omega)^{-1}$ characterizes the magnetoelastic gap at the point of resonance ($\Delta q = 0$) in the absence of losses ($Q_{\mathrm{M}} = Q_{L(\mathrm{T,S})} = 0$ [1], $\delta f_{\mathrm{me}} = \delta\omega_{\mathrm{me}}/\pi$; insert I in Fig. 3).

In a frequency interval where $v \gg v_{\mathrm{S}}$, the group velocity of the elastic waves is usually positive ($V^g_{\mathrm{L(T,S)}} > 0$) and that of the BVMSW is negative ($V^g_{\mathrm{M}} < 0$). The difference in the signs of the group velocities of the BVMSWs and elastic modes causes stop bands [7] in the spectrum to appear as a result of interaction between the waves. The dispersion and damping of MEWs for this case are shown in insert I to Fig. 3. Here, the region where the BVMSW with the number $m = 0$ is in resonance with the longitudinal Lamb mode with the number $l = 6$ (this region corresponds to area C in Fig. 1) is shown. In the absence of dissipation (dashed curves), the dispersion curves of the interacting waves "repulse," forming the stop band $\delta f^r_{\mathrm{me}} \le \delta f_{\mathrm{me}}$ wherein the waves are in synchronism and decay with the maximal decrement at the synchronism frequency [1]:

$$q'' = \delta\omega_{\mathrm{me}}\mathrm{Im}(V^g_{\mathrm{M}}V^g_{\mathrm{L(T,S)}})^{-\frac{1}{2}}. \tag{18}$$

If the wave matching is achieved at frequencies for which the Lamb mode dispersion is negative (crosses in Fig. 1), the group velocities of the waves are codirected and the resonance interaction repulses the dispersion curves, which is typical of this case [1]. This is illustrated by insert II in Fig. 3, where the region of phase matching between the BVMSW mode with the number $m = 0$ and the transverse Lamb mode with the number $t = 11$ is shown (this region corresponds to the area D in Fig. 1).

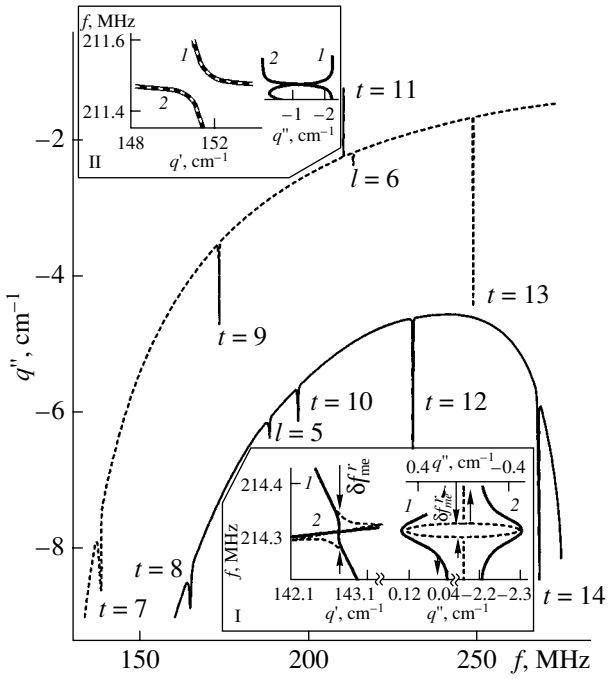


Fig. 3. Frequency dependences of the spatial damping decrement of the BVMSWs $m = 0$ (dashed curves) and $m = 1$ (continuous curves) in the spectral range corresponding to the fast MEWs. The numbers of these modes are indicated alongside with the peaks of hybridization with the elastic modes. The inserts I and II show the dispersion and damping of the MEWs near the resonance of the BVMSW $m = 0$ with the longitudinal Lamb mode $l = 6$ and transverse Lamb mode $t = 11$. Dashed curves, without dissipation; continuous curves, with dissipation. Figures 1 and 2 in the inserts are used to show the one-to-one correspondence between the dispersion, $f(q')$, and loss, $f(q'')$, curves.

For the given dissipation parameters, the repulsion of the dispersion curves in most cases changes to anomalous regions like those shown by the continuous curves in the insert I to Fig. 3. At the same time, the behavior of the dispersion curves near the resonance of the BVMSW mode with the number $m = 0$ with the Lamb mode (number $t = 11$) remains virtually unchanged. It turns out that the domains of anomalous dispersion for this resonance form if the FMR linewidth increases threefold or the elastic viscosity increases 200 times.

Figure 3 shows the frequency dependence of the spatial damping decrement for two first BVMSW modes in the plate. The negative values of the decrement mean that the direction of the BVMSW group velocity is opposite to that of the x axis. It can be seen that the BVMSW damping is the greatest (stop bands arise) at the frequencies corresponding to the resonance with the transverse Lamb modes and at the frequencies corresponding to the resonance with the longitudinal Lamb modes if their frequencies are close enough to the resonances with the transverse modes. The BVMSW damping at the frequencies of resonances with the

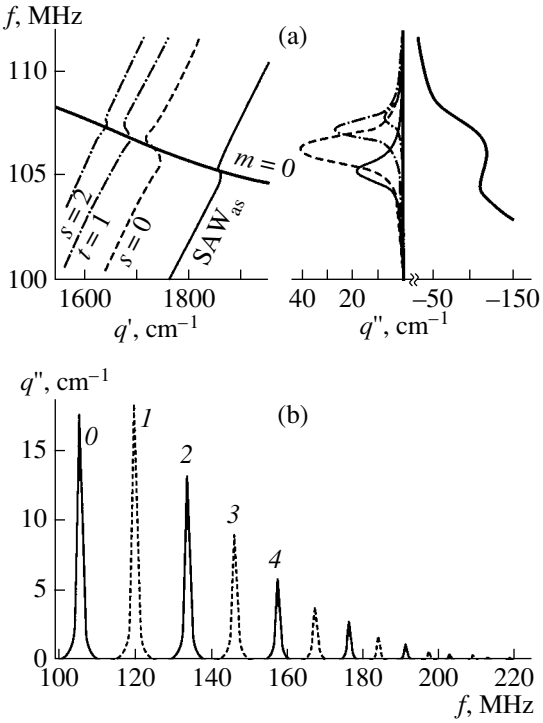


Fig. 4. (a) Dispersion and damping of the MEWs near the resonance of the BVMSW mode $m = 0$ with the antisymmetric surface Lamb mode (SAW_{as}) and transverse Lamb mode $t = 1$ and also with the shear modes $s = 0$ and $s = 2$ (with dissipation); (b) frequency dependences of the spatial damping decrement of the symmetric (dashed curves) and antisymmetric (continuous curves) surface Lamb modes. Near the peaks of hybridization with the BVMSW modes, the mode numbers are indicated.

purely shear modes is one or two orders of magnitude weaker and is not shown in Fig. 3. Such a weak influence of the purely shear waves on the BVMSW damping is because their coupling in equations of motion (1) is of the order of $\sim qb$ and turns to be small in the region of the fast MEWs, where $q \longrightarrow 0$. Generally, in the region where $v \gg v_S$, the BVMSW modes interact most effectively with the transverse Lamb modes.

It should also be mentioned that since the dissipation in the magnetic subsystem is higher than in the elastic one, the interaction with the Lamb mode having a negative slope of the dispersion ($t = 11$) decreases the BVMSW damping.

The resonance interaction at $v \approx v_S$ will be considered by the example of the resonances falling into the area F in Fig. 1. Figure 4 shows the calculated dispersion and attenuation curves for the magnetoelastic waves from this area.[2] Unlike the region of the fast MEWs, here the efficiencies of the BVMSW interaction with the Lamb modes and with the purely shear modes are the same. Note also that in this case, the magnetoelastic gap is $\delta f_{me} \approx 1$ MHz, that is, exceeds the maximum value of δf_{me} from the region of the fast MEWs by almost one order of magnitude.

Losses result in the formation of the clearly defined domains of anomalous dispersion and in the oscillations of the damping decrement for the elastic modes at the resonance frequencies. As an example, Fig. 4b shows the frequency dependences of the damping for the symmetric and antisymmetric surface Lamb modes. The oscillation decays with growing number of the BVMSW mode, which reflects the decrease in the overlap integral by the near-exponential law. This law is violated only for the resonance with the BVMSW mode with the number $m = 0$ presumably because of the effect of the resonances between adjacent volume elastic modes. In the latter case, the approach based on Eqs. (17) is inapplicable since more than two waves take part in interaction.

RESONANCE INTERACTION OF LAMB AND SHEAR ELASTIC MODES

As follows from Eqs. (1), at $b \neq 0$, the shear and Lamb elastic modes are coupled together through the magnetic subsystem. Consider such an interaction under the conditions of wave phase matching for the resonance of the transverse Lamb mode $t = 6$ with the shear mode $s = 7$. This resonance takes place at $f_E \approx 144.8$ MHz and $q_E \approx 865$ cm^{-1} (the area E in Fig. 1) (the calculations show that the interaction of elastic modes with other numbers is qualitatively similar). Equation (9) takes the form

$$\Delta S \Delta L = \delta_3.$$

The conditions for elastic wave matching in the area E are fulfilled at any value of the magnetic field.[3] However, it might be expected that the efficiency of interaction between the elastic waves will vary according to the spacing between the dispersion curves of the magnetostatic waves and the points of intersection between the dispersion curves of the elastic waves. Indeed, the amplitude of the forced oscillations of magnetization, which accompany the propagation of the elastic waves and provide their coupling, directly depends on the phase mismatch between the elastic waves and the BVMSWs.

The insert in Fig. 5 shows the dispersion and the damping decrement for the elastic waves in the area E for the field $H_0 = 125$ Oe. With this field, the lower (short-wave) boundary of the region of BVMSW existence is $f_H = 350$ MHz, which markedly exceeds the resonance frequency f_E. In the absence of losses, the gap $\delta f_e \sim b^2$ and equals $\delta f_e \approx 1$ kHz in this case. For the given

[2] In Fig. 4a, only those elastic modes obeying the selection rule for the interaction with the BVMSW mode with the number $m = 0$ are shown.

[3] The position of the point of intersection (f_E, q_E) between the dispersion curves of the elastic modes depends on the magnetic field because of the change in the renormalization of the velocity of sound.

parameters of relaxation in the elastic and magnetic subsystems of the ferromagnetic, the gap in the area E at $H_0 = 125$ Oe is retained unlike the interaction of the elastic modes with the BVMSWs (insert II in Fig. 3; here, the dissipation and resonance interaction affect the spectrum to the same extent). Note that the gap in the frequency dependences of the elastic wave damping decrements in the resonance region [1] is lacking (the dashed curves in the insert to Fig. 5). Such behavior of the dispersion and damping of the waves in the vicinity of the resonance corresponds to the case of "small" losses [1]:

$$\delta\omega_e^2 \gg (Q_S - Q_T)^2 V_T^g V_S^g.$$

Generally, the resonance interaction for the field $H_0 = 125$ Oe affects the dispersion and damping of the elastic waves in the same manner as the resonances of the magnetostatic waves with the elastic [1] or exchange [8] ones. However, if the area E is close to the region of BVMSW existence, the frequency dependences of the elastic wave damping decrements $q''(\omega)$ in the resonance region are markedly different (Fig. 5). In the narrow frequency interval near the resonance frequency f_E, the frequency dependences of the decrements not only intersect but also move apart. It is essential that the repulsion of the dispersion curves for the shear and Lamb modes is retained and has the form similar to that shown in the insert to Fig. 5a. The additions to the decrements $\delta q''$ (Fig. 5) do not exceed the extra damping caused by the coupling with the magnetic subsystem.

The run of the curves $q''(\omega)$ near the resonance of the shear and Lamb elastic waves (Fig. 5) can be explained with expression (10) after relevant modifications. It is necessary to take into account the contribution of relaxation processes to the wave coupling δ_3 [9]:

$$\tilde{\delta}_3 \Rightarrow \delta_3\left(1 - i\frac{\gamma\Delta H}{\omega_H}\Phi\right),$$

where Φ is some function of ω and q of the order of unity. The additions due to elastic viscosity are omitted, because they are small in our ranges of frequencies and other parameters.

If the elastic wave decrements are so close that $|Q_S - Q_T| \ll \sqrt{\delta\omega_e^2/V_S^g V_T^g}(\gamma\Delta H/\omega_H)\Phi$ (curves in Fig. 5 for $H_0 = 55$ Oe), the expression for $\delta q''$ is greatly simplified:

$$\delta q'' \approx \mp\frac{1}{2}\sqrt{\frac{\delta\omega_e^2}{V_S^g V_T^g}}\frac{\gamma\Delta H}{\omega_H}\Phi.$$

The repulsion of the elastic wave decrements at the resonance frequency can qualitatively be related to the variation of the elastic wave coupling with the magnetic subsystem and, hence, of the insertion losses. Obviously, this effect will be the most pronounced, i.e., $\delta q'' \sim Q_{S,T}$, if the contribution of the magnetic losses to the damping of the elastic modes far exceeds that from the intrinsic viscous losses.

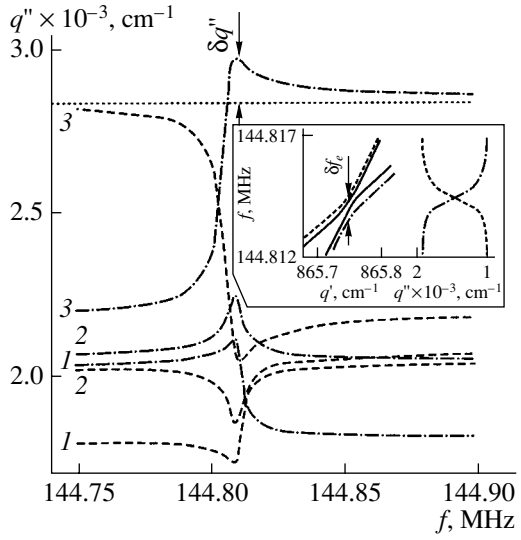


Fig. 5. Frequency dependences of the elastic wave decrements near the resonance of the elastic mode $s = 7$ with the transverse Lamb mode $t = 6$. H_0 = (*1*) 60, (*2*) 55, and (*3*) 45 Oe. The insert shows the dispersion and damping of these modes at $H_0 = 125$ Oe (continuous curves, without dissipation; dashed and dash-and-dot lines, with dissipation).

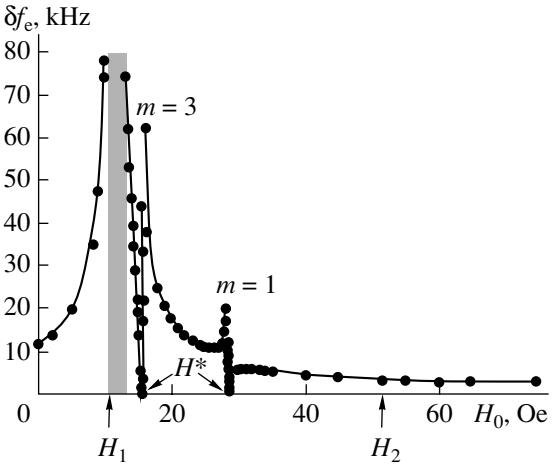


Fig. 6. Split of the dispersion curves for the elastic modes, δf_e, at the point of the synchronism of the shear mode $s = 7$ with the transverse Lamb mode $t = 6$ vs. magnetic field in the absence of dissipation.

Figure 6 shows the dependence of the gap δf_e on the magnetic field. The value of δf_e appreciably increases in the field interval $H_1 - H_2$, where BVMSWs exist. At $H_0 = 0$, the gap is finite, whereas $\delta f_e \longrightarrow 0$ as the field grows. Within the field interval $H_1 - H_2$, there is a number of values at which the BVMSW modes also find

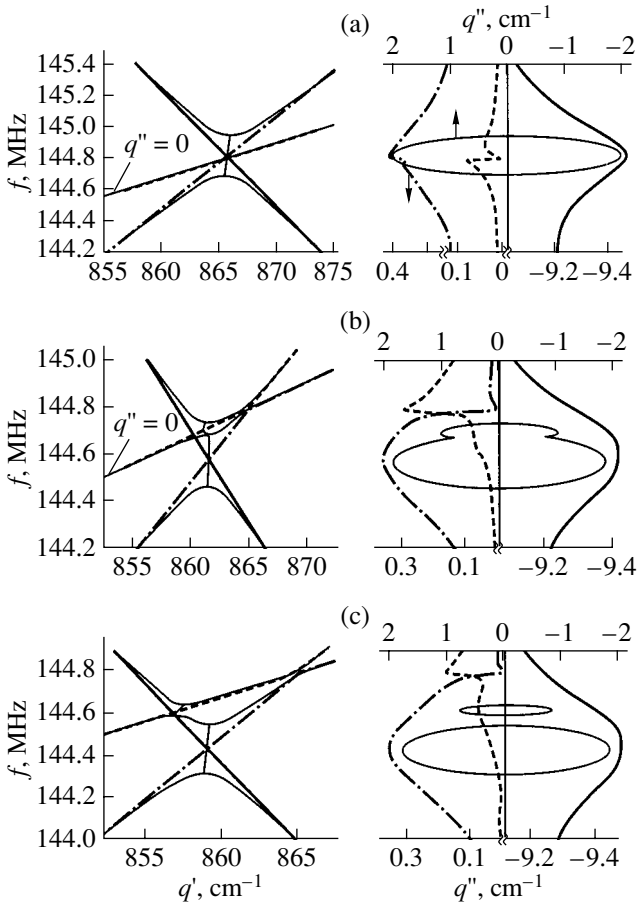


Fig. 7. Dispersion and damping of the MEWs near the resonance of the BVMSW mode $m = 1$ with the transverse Lamb mode $t = 6$ and with the shear mode $s = 7$. H_0 = (a) 28.259, (b) 28.1, and (c) 28 Oe. Thin continuous lines, without dissipation; thick lines, with dissipation.

themselves in the region of elastic wave resonance, so that three waves resonate (see the points of discontinuities).[4] In the vicinity of the triple resonances, two effects are observed. First, the efficiency of the elastic mode interaction increases (δf_e are close to the values of δf_{me} from the interval of the fast MEWs), which is a consequence of the increased sensitivity of the magnetic subsystem to external excitations. Second, the elastic resonances are suppressed (see the values of the fields H^*). The latter effect is apparently related to the competition between the dispersion modifications due to the interaction of the elastic modes with each other, on the one hand, and to the interaction of the elastic modes with the BVMSWs, on the other.

In order to explain the behavior of the waves under the conditions of triple resonance, we will turn to Fig. 7, where the dispersion (left column) and the damping (right column) of the waves in the area E are shown for the fields H_0 at which the BVMSW mode with the number $m = 1$ falls into the area E. The calculations performed without regard for the dissipation in the spin and elastic systems are shown by thin continuous lines. The dispersion laws and the frequency dependences of the imaginary parts of the wave numbers, q'', calculated with regard to the losses are shown by the thick curve for the BVMSW mode and by the dashed and dash-and-dot curves for the elastic modes.

Figure 7a demonstrates the case when the dispersion curves intersect practically in one point, i.e., triple resonance. It can be seen that if dissipation is neglected, the resonance interaction shows up only in the dispersion and damping laws for the BVMSW and Lamb modes, which behave similarly to the case shown in the insert I in Fig. 3. The position of the shear mode in the plane (Ref, Req) remains nearly unchanged and the gap is $\delta f_e \approx \delta f_{me}/2 \approx 100$ kHz, where δf_{me} characterizes the repulsion of the dispersion curves for the BVMSW and the Lamb mode at the resonance frequency.

The aforesaid can be easily understood if it is considered that the resonance E takes place in the region of the fast MEWs, where the coupling of the BVMSW with the pure shear modes is at least one order of magnitude weaker than with the Lamb modes. Therefore, the hierarchy of the interactions is such that the BVMSW and Lamb modes interact in the first place and form the stop band, where they experience losses given by (11). The result of this process is the transition of the dispersion curves from the real plane (Ref, Req) to the complex space (Ref, Req, Imq). Taking into account the coupling with the shear mode [the term with δ_2 in (9)] alter the position of the dispersion curves insignificantly, since the dispersion curve of the shear mode is offset from the point of resonance between the BVMSW and the Lamb mode by the value of $\Delta\omega = 2\pi\delta f_e \gg \sqrt{\delta_2}, \sqrt{\delta_3}$.

The calculations performed with regard for the given relaxation parameters show that the dispersion laws almost coincide with those when the resonance interaction is absent, whereas the frequency dependences of the decrements exhibit anomalous domains caused by the interaction of the elastic modes both with the BVMSW and with each other.

Figures 7b and 7c illustrate the dispersion and damping laws when the BVMSW mode leaves the region of elastic wave resonance under the action of the magnetic field. As long as the point E, where the dispersion curves of the elastic modes intersect, is within the stop band formed by the interaction of the BVMSW and Lamb modes, the dispersion curves and the damping of the waves near the resonance are similar to those shown in Fig. 7a. For fields H_0 at which the frequency of the resonance E is beyond the stop band because of the resonance of the BVMSW with the Lamb mode, the triple resonance splits into three double resonances (Fig. 7b). In this case, the resonances of the BVMSW

[4] For the shaded region of the field interval, where the density of the BVMSW modes is high and the elastic modes interact with several of them, the calculations were not carried out.

with the Lamb and shear elastic modes overlap to form the integral stop band in the spectrum and the laws of dispersion and damping for the elastic waves near the resonance *E* become similar to those shown in Fig. 5. At H_0 = 28 Oe, the BVMSW mode leaves the region of elastic wave resonance and insignificantly affects the resonance interaction of the elastic waves (Fig. 7c).

WAVE INTERACTION AT STRONG MAGNETOELASTIC COUPLING

In most cases discussed above, the waves interacted under conditions of weak magnetoelastic coupling, when the repulsions (δf_e and δf_{me}) of the dispersion curves at the resonance frequency are considerably smaller than the frequency interval Δf between the adjacent modes in the spectrum: $\Delta f \gg \delta f_{me}, \delta f_e$.

However, at certain relationships between the parameters of the problem (wave frequencies, thickness and magnetoelastic coupling constant of the ferrite layer, and magnetic field), multiwave resonances may take part in the formation of the MEW spectrum in layered ferrite structures, when the conditions of strong magnetoelastic coupling, $\Delta f \le \delta f_{me}, \delta f_e$, are fulfilled.

An example of the multiwave resonance is the above triple resonance involving the BVMSW mode with the number m = 1 and the shear and transverse Lamb modes with s = 7 and t = 6. However, such a resonance has a random character; namely, it appears only at a certain value of the magnetic field and at a minor change in the field, ~1 Oe, the BVMSW mode drops out of synchronism. It should also be noted that the low efficiency of interaction of the BVMSW with the shear mode allows us to consider the participation of the latter as a weak perturbation with relation to the strong interaction of the BVMSW with the Lamb mode (Fig. 7).

Aside from the triple resonances considered, other resonances are possible in our case. They are characterized, first, by a greater number of participating waves and, second, by comparable efficiencies of the interaction of all elastic waves with the BVMSW. An example is the resonance interaction of the BVMSW with the elastic waves in the range $v \approx v_S$, where the high-frequency crowding of the elastic modes occurs (the area *F* in Fig. 1).

Consider the dispersion and damping of the waves under strong magnetoelastic coupling by the example of the resonance of the fundamental BVMSW mode with the antisymmetric surface Lamb mode (SAW_{as}), transverse Lamb mode t = 1, and shear modes s = 0 and s = 2 (Fig. 8). The resonance interaction of these waves in the case of weak coupling is shown in Fig. 4a and corresponds to the magnetoelastic coupling constant b = 1600. As b increases, the branches of the spectrum approach each other in the anomalous dispersion region and intersect at a certain critical value. In this case, each of the dispersion curves on either side of the intersection point corresponds to different elastic modes. The

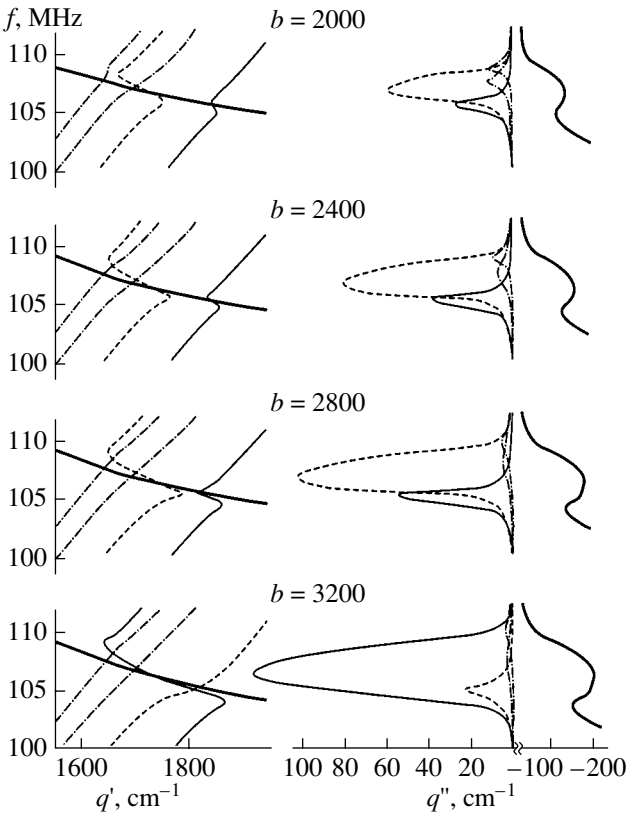


Fig. 8. Dispersion and damping of the MEWs near the resonance of the BVMSW mode m = 0 with the antisymmetric surface Lamb mode (SAW_{as}) and the transverse Lamb mode t = 1 and also with the shear modes s = 0 and s = 2 with dissipation at different values of the magnetostriction constant.

attenuation curves are qualitatively modified; namely, a gap forms at the point of their intersection. The damping of one branch becomes much lower than in the case $\Delta f_e > \delta f_{me}$, which can be interpreted as falling out of resonance with the BVMSW. The feature of the multiwave resonances in this region is that they may involve a great number of the elastic modes. The overall magnetoelastic resonance covers more than two elastic modes by a chain mechanism, successively involving the nearest neighbors. In this case, the domain of anomalous dispersion, to which the maximal damping corresponds, originates from the mode that has the lowest phase velocity and ends at the mode with the highest phase velocity. The dispersions of the other modes smoothly occupy the place of the neighbors with a lesser phase velocity, and their damping peaks markedly decrease as against the case $\Delta f_e > \delta f_{me}$. For the magnetoelastic interaction constant b = 3200, the surface mode turns out to be involved in the resonance interaction of the BVMSW and volume elastic modes. In this case, the surface mode transforms into the volume one.

CONCLUSIONS

Thus, the results obtained are as follows.

(1) Magnetostriction splits the cutoff frequencies both of the transverse Lamb and shear modes and of the BVMSW modes. This appears as the offsets of the cutoff frequencies of the transverse Lamb modes and the BVMSW modes from the initial position and also in the resonance interaction between the BVMSW modes of equal evenness.

(2) BVMSWs can interact with both the Lamb and the shear modes of the plate. This usually shows up in the formation of the stop bands δf^r_{me} and in the oscillations of the damping at the frequencies of wave matching. In the range $v \approx v_S$, the efficiencies of the resonance interaction of all elastic modes with the BVMSWs turn out to be of the same order. In the region of fast magnetoelastic waves ($v \gg v_S$), the BVMSWs interact most efficiently with the transverse Lamb modes, which partially indicates the changes in the cutoff frequencies of these modes.

(3) Magnetostriction results in the resonance interaction of the shear and elastic modes of the plate. Here, the parameter δf_e, first, is proportional to the square of the magnetoelastic coupling constant (unlike the resonances involving BVMSWs) and, second, may serve as a measure of the repulsion of both the dispersion and damping curves.

(4) In the range $v \approx v_S$, where the crowding of the dispersion curves for the elastic modes takes place, several elastic modes of the plate resonate with the BVMSW mode. The feature of such multiwave resonances is that the surface modes may transform into the volume ones and the shear modes, into the Lamb modes or into shear modes with other numbers.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 99-02-17660 and 01-02-06108), CDRF (grant no. REC-006), and the Federal Program “Integration” (grant no. 696.3).

REFERENCES

1. Yu. V. Gulyaev and P. E. Zil’berman, Izv. Vyssh. Uchebn. Zaved., Fiz. **31** (11), 6 (1988).
2. B. N. Filippov and V. P. Lukomskiĭ, Fiz. Met. Metalloved. **34**, 682 (1972).
3. R. E. Camley, J. Appl. Phys. **50**, 5272 (1979).
4. R. W. Damon and J. R. Eshbach, J. Phys. Chem. Solids **19**, 308 (1961).
5. I. A. Viktorov, *Surface Acoustic Waves in Solids* (Nauka, Moscow, 1981).
6. R. C. Le-Crow and R. L. Comstock, in *Physical Acoustics: Principles and Methods*, Ed. by W. P. Mason (Academic, New York, 1965; Mir, Moscow, 1968), Vol. 3, Part B.
7. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
8. A. V. Lugovskoĭ and P. E. Zil’berman, Fiz. Tverd. Tela (Leningrad) **24**, 458 (1982) [Sov. Phys. Solid State **24**, 259 (1982)].
9. A. I. Akhiezer, V. G. Bar’yakhtar, and S. V. Peletminskiĭ, *Spin Waves* (Nauka, Moscow, 1967; North-Holland, Amsterdam, 1968).

Translated by M. Lebedev

Technical Physics, Vol. 47, No. 1, 2002, pp. 49–57. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 51–59.

SOLIDS

Effective Parameters of Multicomponent Dielectrics with Hexagonal Structure

Yu. P. Emets

Institute of Electrodynamics, National Academy of Sciences of Ukraine, Kiev, 03680 Ukraine

e-mail: emets@irpen.kiev.ua

Received March 29, 2001

Abstract—The effective parameters of a matrix dielectric medium that contains three types of unidirectional cylindrical inclusions forming the hexagonal lattice are calculated. The electric field and the average characteristics of the system are analytically calculable if the concentration of the inclusions is low. Small parameters characterizing the system are indicated, and the general form of the expression for the effective permittivity of anisotropic multicomponent media is presented. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In this work, we study the electrical properties of inhomogeneous dielectric media with the symmetric packing of dissimilar unidirectional cylindrical fibers. This problem is frequently encountered in many fields of physics and biology, as well as in applications [1–5].

There are five possible ways of packing similar fibers, depending on the number of polygons continuously filling the plane. These are known to be parallelograms, rectangles, triangles, squares, and hexagons with the respective axes of one-, two-, three-, four-, and six-fold rotational symmetry. The fibers are placed in the cells of a 2D mesh. Among these packings, the hexagonal texture (lattice) is of greatest interest. In this case, the composite material retains the periodic texture when inclusions of three sorts are symmetrically packed in the matrix. The medium under study is thus four-component. The arrangement of dissimilar inclusions in the system is akin to the presence of color symmetry, where identically colored inclusions are assumed to have certain physical properties.

The properties of two-component hexagonal structures consisting of the matrix and identical parallel fibers were studied in [6]. A more general problem (as applied to the analysis of thermal conduction) was considered in [7]. In both papers, the Rayleigh methods [8] were employed. Using the classical potential theory, Rayleigh has developed efficient methods for studying two- and three-component matrix media with periodically arranged, respectively, cylindrical fibers and spherical inclusions.

However, in the systems with dissimilar fibers of various radii and physical properties, as in our system, the Rayleigh methods of calculating the electric field fail. In this case, one should use another computing scheme that is based on the summation of the fields generated by interacting inclusions and uses the exact solution of the problem of two interacting parallel cylinders [9].

Although our calculations directly concern dielectric media, the mathematical approach used can be extended for calculating transport coefficients in the theory of elasticity, fluid dynamics, magnetostatics, and in the theory of thermal and electrical conduction.

LOCAL ELECTRIC FIELD

Consider a dielectric medium that consists of the matrix with a permittivity ε_1 and symmetrically arranged unidirectional parallel cylindrical fibers of three sorts. The fibers with permittivities ε_2, ε_3, and ε_4 and radii r_1, r_2, and r_3 alternate, forming the hexagonal structure. A fragment of the infinite medium is shown in Fig. 1. The applied electric field $\mathbf{E}_0$ lies in the plane *xy* normally to the axes of the fibers. For such a geometry, the equations of electrostatics,

$$\nabla \cdot \mathbf{D} = 0, \quad \nabla \times \mathbf{E} = 0, \quad \mathbf{D} = \varepsilon \mathbf{E}, \qquad (1)$$

where **D** and **E** are the induction and the strength of the electric field, are two-dimensional and coincide with the Cauchy–Riemann conditions. This allows us to pass into the plane of the complex variable *z* and introduce the complex functions of the electric field

$$D(z) = D_x - iD_y, \quad E(z) = E_x - iE_y \quad (z = x + iy). \qquad (2)$$

Because of the symmetrical structure of the medium, the electric field is periodical. Therefore, it would suffice to calculate the field in one structure element, i.e., in one hexagonal cell with one inclusion (Fig. 1). The field inside and in the vicinity of the inclusion separated depends on the fields of other inclusions. The net effect can be considered as the interaction of the inclusion separated with each of the remaining inclusions. When solving this general problem, one can consider the subproblem of finding the electric field of

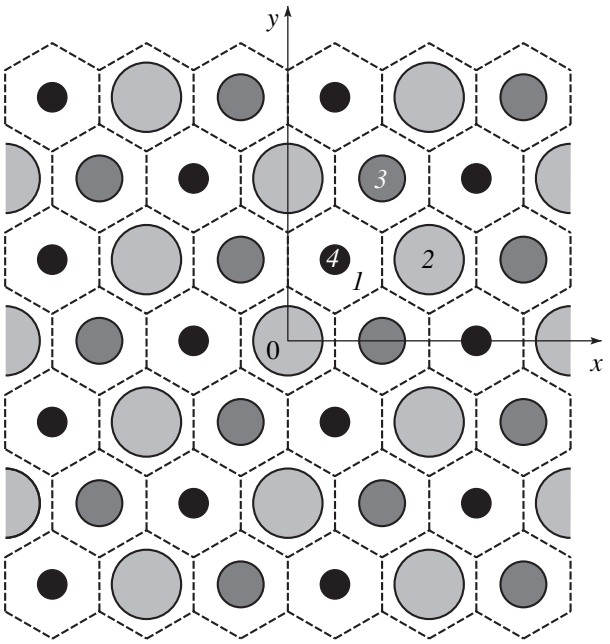


Fig. 1. Fragment of a four-component (*1–4*) dielectric material with the hexagonal packing of unidirectional fibers.

two parallel dielectric cylinders dipped into a dielectric medium under a uniform external field. This subproblem has an exact solution in the general case, when the radii and the permittivities of the cylindrical bodies differ and the bodies are arbitrarily arranged [9].

Mathematically, the interplay of the cylinders is expressed by dipole–dipole interaction. These are linear induced dipoles. The procedure for calculating the dipole coordinates and moments was outlined in [9]. In this work, we study a system with a small concentration of inclusions; hence, the single-dipole approximation of the electric field is valid.

To be definite, we assume that the origin coincides with the center of the inclusion of permittivity ε_2. Then, the electric field in the surrounding hexagon is given by

$$E_1(z) = E_0 - \bar{E}_0 \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} [\Delta_{12} r_1^2 (z - a_{mn})^{-2} + \Delta_{13} r_1^2 (z - b_{mn})^{-2} + \Delta_{14} r_1^2 (z - c_{mn})^{-2}],$$

$$E_2(z) = (1 + \Delta_{12})\left\{E_0 + \bar{E}_0\left[\Delta_{12} r_1^2 z^{-2} - \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} [\Delta_{12} r_1^2 (z - a_{mn})^{-2} + \Delta_{13} r_1^2 (z - b_{mn})^{-2} + \Delta_{14} r_1^2 (z - c_{mn})^{-2}]\right]\right\}. \quad (3)$$

Here, $E_1(z)$ and $E_2(z)$ are the complex strengths of the electric fields in the matrix and in the inclusion, respectively, and $E_0 = E_{0x} - iE_{0y}$ is the uniform applied electric field (the bar over E_0 means complex conjugation). The dimensionless parameters Δ_{12}, Δ_{13}, and Δ_{14} are expressed as

$$\Delta_{12} = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2}, \quad \Delta_{13} = \frac{\varepsilon_1 - \varepsilon_3}{\varepsilon_1 + \varepsilon_3}, \quad \Delta_{14} = \frac{\varepsilon_1 - \varepsilon_4}{\varepsilon_1 + \varepsilon_4} \quad (-1 \le \Delta_{12}, \Delta_{13}, \Delta_{14} \le 1). \quad (4)$$

The electric field is seen to be the sum of the uniform field and the field produced by the infinite number of induced dipoles whose coordinates a_{mn}, b_{mn}, and c_{mn} coincide with the centers of the inclusions:

$$a_{mn} = \left\{h(3m + i\sqrt{3}n), h\left[3m + \frac{3}{2} + i\frac{\sqrt{3}}{2}(2n + 1)\right]\right\},$$

$$b_{mn} = \left\{h(3m + 1 + i\sqrt{3}n), h\left[3m + \frac{5}{2} + i\frac{\sqrt{3}}{2}(2n + 1)\right]\right\}, \quad (5)$$

$$c_{mn} = \left\{h(3m + 2 + i\sqrt{3}n), h\left[3m + \frac{1}{2} + i\frac{\sqrt{3}}{2}(2n + 1)\right]\right\},$$

where h is the linear size of the period along the x axis and $m, n = 0, \pm1, \pm2, \ldots$ (Fig. 1).

With the origin placed at the center of the inclusion with the permittivity ε_3, expressions (3) hold if the circular permutation of the parameters is made: $\Delta_{12} \longrightarrow \Delta_{13}$, $\Delta_{13} \longrightarrow \Delta_{14}$, and $\Delta_{14} \longrightarrow \Delta_{12}$; $r_1 \longrightarrow r_2$, $r_2 \longrightarrow r_3$, and $r_3 \longrightarrow r_1$. Similarly, for the inclusion with the permittivity ε_4: $\Delta_{12} \longrightarrow \Delta_{14}$, $\Delta_{13} \longrightarrow \Delta_{12}$, and $\Delta_{14} \longrightarrow \Delta_{13}$; $r_1 \longrightarrow r_3$, $r_2 \longrightarrow r_1$, and $r_3 \longrightarrow r_2$.

The interaction between the inclusions varies in inverse proportion to the center-to-center distance squared according to the properties of a linear dipole field. Therefore, in (3), the terms with larger subscripts m and n (or at least with one of them) are usually rejected in practical calculations. This simplification allows the use of finite sums in place of infinite ones.

Expressions (3) are written in the single-dipole approximation, where only the first induced dipoles are considered. Their moments are proportional to the parameters $\Delta_{1\vartheta}$ ($\vartheta = 2, 3, 4$) in the first and second power. The moments are also proportional to the squares of the relative radii r_j/h ($j = 1, 2, 3$) or, what is the same, to the cross-sectional area of the fibers ($s_j = \pi r_j^2/h^2$). Subsequent approximations will contain dipoles whose moments are proportional to higher powers of $\Delta_{1\vartheta}$ and s_j. These parameters enter the dipole moments multiplicatively, so that expressions (3) are applicable if at least one of the parameters is small.

If the interaction between the inclusions is ignored, the infinite sums in (3) can be omitted. In this case, the field inside the inclusions is uniform, while that outside the inclusions is represented by a dipole. Such an approximation is used for calculating the effective parameters of weakly inhomogeneous materials [10, 11].

LOCAL FIELD AVERAGING

The macroscopic properties of an inhomogeneous dielectric are defined by the averaged material equation

$$\langle \mathbf{D} \rangle = \hat{\varepsilon}_{\text{eff}} \langle \mathbf{E} \rangle, \quad (6)$$

where the angular brackets refer to volume-averaged quantities. Here, averaging is made over the unit cell of the double-periodic system.

On the average, the composite under study can be considered as anisotropic because of the hexagonal packing of unidirectional cylindrical fibers. In the plane normal to the cylinder axis, the macroscopic properties of the material are described by the effective permittivity tensor $\hat{\varepsilon}_{\text{eff}}$. The components of the symmetric tensor $\hat{\varepsilon}_{\text{eff}}$ reduced to the principal axes, $\hat{\varepsilon}_{\text{eff}} = \{\varepsilon_{\text{eff}\,xx}, \varepsilon_{\text{eff}\,yy}\}$ are given by

$$\langle D \rangle_x = \varepsilon_{\text{eff}\,xx} \langle E \rangle_x, \quad \langle D \rangle_y = \varepsilon_{\text{eff}\,yy} \langle E \rangle_y. \quad (7)$$

The averaged fields are calculated by averaging expressions (3) over the unit cell of the period of the system. As a unit cell, we can take the equilateral triangle *ABC* (Fig. 2). If the external field is aligned with the *x* axis ($E_0 = E_{0x}$), this axis will be a line of force of the field, because the system is mirror-symmetric about the real axis. Force lines are also all lines that are parallel to the real axis and are offset from it by a distance multiple of the length $BD = \sqrt{3}\,h/2$ (the height of the triangular cell). Thus, the triangle *ABC* lies between two adjacent lines of force $y = 0$ and $y = \sqrt{3}\,h/2$, the side *AC* being coincident with the line $y = 0$ and the vertex *B* lying in the other line, $y = \sqrt{3}\,h/2$. To such a cell, the inclusion of each of the types contributes 1/6 of its plane section.

As usual, the averaging of the field over the cell can be replaced by taking the contour integrals. Let the external field be applied along the *x* axis ($E_0 = E_{0x}$). Then,

$$\langle E \rangle_x = \frac{1}{h}\int_A^C \mathrm{Re}E(x)dx, \quad \langle D \rangle_x = \frac{2}{\sqrt{3}h}\int_D^B \mathrm{Re}E(z)dz. \quad (8)$$

If the field is aligned with the *y* axis ($E_0 = -iE_{0y}$),

$$\langle E \rangle_y = \frac{2}{\sqrt{3}h}\int_D^B \mathrm{Im}E(z)dz, \quad \langle D \rangle_y = \frac{1}{h}\int_A^C \mathrm{Im}E(x)dx. \quad (9)$$

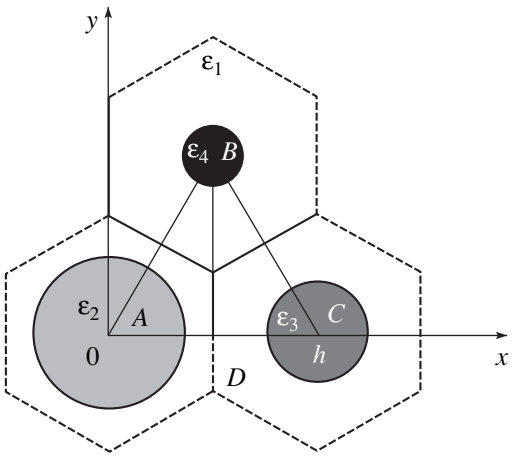


Fig. 2. Unit cell of the periodic dielectric material that is used for field averaging.

In the latter case, lines that previously were lines of force become equipotential lines. The simple yet demanding some accuracy calculation of the first integral in (8) yields for the averaged field in the cell along the *x* axis

$$\begin{aligned}\langle E \rangle_x &= 1 + q(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2) \\ &+ \Delta_{12}\Delta_{13}[r_1^2 M(r_2) + r_2^2 M(r_1)] + \Delta_{12}\Delta_{14}r_3^2 N(r_1) \\ &+ \Delta_{13}\Delta_{14}r_3^2 N(r_2) + \frac{3}{2}[\Delta_{12}^2 V(r_1) + \Delta_{13}^2 V(r_2)].\end{aligned} \quad (10)$$

Expression (10) and all subsequent formulas for the averaged electric field are written in the relative values

$$\langle E \rangle_x^* = \langle E \rangle_x / |E_0|, \quad r_j^* = r_j/h \ (0 \le r_j^* < \sqrt{3}/2); \quad j = 1, 2, 3.$$

In what follows, the asterisks will be omitted for the sake of simplicity. In (10), *q* is the numerical parameter whose exact value is defined by two formulas, one of which is given in Appendix A. The calculation yields $q = 0.8061$. The expressions for the functions $M(\ldots)$, $N(\ldots)$, and $V(\ldots)$ are also derived in Appendix A. This appendix summarizes all auxiliary formulas necessary for using expression (10) and subsequent expressions (11)–(13).

According to (10), the average electric field $\langle E \rangle_x$ is asymmetric with respect to the parameters Δ_{12}, Δ_{13}, and Δ_{14}, which are the reduced permittivities of the inclusions. This theoretical finding is inconsistent with what would be expected for the structurally symmetric material. Note, however, that the form of the expression for $\langle E \rangle_x$ depends, in general, on the position of the origin. In our case, the origin is at the center of the inclusion with the permittivity ε_2 (Figs. 1, 2). Obviously, the position of the origin has no effect on the absolute value of $\langle E \rangle_x$.

For example, with the origin placed at the center of the inclusion with ε_3, calculations similar to those used in deriving (10) yield

$$\begin{aligned}\langle E\rangle_x &= 1 + q(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{13}\Delta_{14}[r_2^2M(r_3) + r_3^2M(r_2)] + \Delta_{12}\Delta_{13}r_1^2N(r_2)\\ &+ \Delta_{12}\Delta_{14}r_1^2N(r_3) + \frac{3}{2}[\Delta_{13}^2V(r_2) + \Delta_{14}^2V(r_3)].\end{aligned} \quad (11)$$

Expression (11) is also asymmetric with respect to Δ_{12}, Δ_{13}, and Δ_{14}, which is most evident when (10) and (11) are compared.

Finally, with the origin placed at the center of the inclusion with ε_4, we have

$$\begin{aligned}\langle E\rangle_x &= 1 + q(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{14}[r_1^2M(r_3) + r_3^2M(r_1)] + \Delta_{12}\Delta_{13}r_2^2N(r_1)\\ &+ \Delta_{13}\Delta_{14}r_2^2N(r_3) + \frac{3}{2}[\Delta_{12}^2V(r_1) + \Delta_{14}^2V(r_3)].\end{aligned} \quad (12)$$

Expressions (10)–(12), while differing in form, give the same value of the average electric field in the system. To make sure of this, one must calculate the field for specific values of the parameters involved.

In order to obtain an expression for $\langle E\rangle_x$ that is symmetric with respect to Δ_{12}, Δ_{13}, and Δ_{14}, it is necessary to add up all three expressions (10)–(12) and to divide the sum by three. Eventually, we come to

$$\begin{aligned}\langle E\rangle_x &= 1 + q(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{13}U(r_1, r_3) + \Delta_{12}\Delta_{14}U(r_1, r_2)\\ &+ \Delta_{13}\Delta_{14}U(r_2, r_3) + \Delta_{12}^2V(r_1)\\ &+ \Delta_{13}^2V(r_2) + \Delta_{14}^2V(r_3),\end{aligned} \quad (13)$$

where

$$U(r_\mu, r_\nu) = \frac{1}{3}\{r_\mu^2[M(r_\nu) + N(r_\nu)] + r_\nu^2[M(r_\mu) + N(r_\mu)]\} \quad (\mu, \nu = 1, 2, 3;\ \mu \neq \nu).$$

Now the field $\langle E\rangle_x$ is symmetric with respect to the above parameters, which is important for the subsequent derivation of the effective permittivity tensor.

The average induction $\langle D\rangle_x$ of the electric field is obtained by taking the second integral in (8). The form of $\langle D\rangle_x$ also depends on the position of the origin. With the origin of Cartesian coordinates fixed at the center of the inclusion with ε_2,

$$\begin{aligned}\langle D\rangle_x &= \varepsilon_1\{1 - p(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{14}r_1^2L(r_3) + \Delta_{13}\Delta_{14}r_2^2L(r_3) + 2\Delta_{14}^2H(r_3)\}.\end{aligned} \quad (14)$$

Here, p is the numerical parameter the expression for which is given in Appendix B. It turns out that $p = 2q$, where q is defined above (after expression (10)). The functions $L(r)$ and $H(r)$, being auxiliary parameters, are given in Appendix B.

If the origin is placed at the center of the inclusion with the permittivity ε_3, we, instead of (14), come to

$$\begin{aligned}\langle D\rangle_x &= \varepsilon_1\{1 - p(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{13}r_2^2L(r_1) + \Delta_{12}\Delta_{14}r_3^2L(r_1) + 3\Delta_{12}^2H(r_1)\}.\end{aligned} \quad (15)$$

For the origin at the center of the inclusion with ε_4, we have

$$\begin{aligned}\langle D\rangle_x &= \varepsilon_1\{1 - p(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{13}r_1^2L(r_2) + \Delta_{13}\Delta_{14}r_3^2L(r_3) + 3\Delta_{13}^2H(r_2)\}.\end{aligned} \quad (16)$$

Expressions (14)–(16) give the same value of $\langle D\rangle_x$ and differ only in form. They are asymmetric with respect to Δ_{12}, Δ_{13}, and Δ_{14}. To obtain a symmetric expression for $\langle D\rangle_x$, it is necessary to add up expressions (14)–(16) and divide the sum by three. This yields

$$\begin{aligned}\langle D\rangle_x &= \varepsilon_1\{1 - p(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{13}G(r_1, r_2) + \Delta_{12}\Delta_{14}G(r_1, r_3)\\ &+ \Delta_{13}\Delta_{14}G(r_2, r_3) + \Delta_{12}^2H(r_1)\\ &+ \Delta_{13}^2H(r_2) + \Delta_{14}^2H(r_3)\},\end{aligned} \quad (17)$$

where

$$G(r_\mu, r_\nu) = \frac{1}{3}[r_\mu^2L(r_\nu) + r_\nu^2L(r_\mu)] \quad (\mu, \nu = 1, 2, 3;\quad \mu \neq \nu).$$

Expressions (13) and (17) define, respectively, the average field strength and induction in the system when the applied field is aligned with the x axis ($E_0 = E_{0x}$). For the external field applied along the y axis ($E_0 = -iE_{0y}$), the average values of the strength and induction are obtained by taking the integrals in (9). Having performed similar calculations, we eventually arrive at

$$\begin{aligned}\langle E\rangle_y &= 1 + p(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{13}G(r_1, r_2) + \Delta_{12}\Delta_{14}G(r_1, r_3)\\ &+ \Delta_{13}\Delta_{14}G(r_2, r_3)\Delta_{12}^2H(r_1)\\ &+ \Delta_{13}^2H(r_2) + \Delta_{14}^2H(r_3),\\ \langle D\rangle_y &= \varepsilon_1\{1 - q(\Delta_{12}r_1^2 + \Delta_{13}r_2^2 + \Delta_{14}r_3^2)\\ &+ \Delta_{12}\Delta_{13}U(r_1, r_2) + \Delta_{12}\Delta_{14}U(r_1, r_3)\\ &+ \Delta_{13}\Delta_{14}U(r_2, r_3) + \Delta_{12}^2V(r_1)\\ &+ \Delta_{13}^2V(r_2) + \Delta_{14}^2V(r_3)\}.\end{aligned} \quad (18)$$

Expressions (18) are symmetric with respect to Δ_{12}, Δ_{13}, and Δ_{14}.

SYMMETRY TRANSFORMATIONS

Comparing expressions (13), (17), and (18), one can ensure that the components of the average electric field in the system satisfy the symmetry relations

$$\begin{aligned}\langle D(\Delta_{12}, \Delta_{13}, \Delta_{14})\rangle_x &= \varepsilon_1\langle E(\Delta_{21}, \Delta_{31}, \Delta_{41})\rangle_y,\\ \langle D(\Delta_{12}, \Delta_{13}, \Delta_{14})\rangle_y &= \varepsilon_1\langle E(\Delta_{21}, \Delta_{31}, \Delta_{41})\rangle_x\end{aligned} \quad (19)$$

or, what is the same,

$$\begin{aligned}\langle D(\Delta_{21}, \Delta_{31}, \Delta_{41})\rangle_y &= \varepsilon_1\langle E(\Delta_{12}, \Delta_{13}, \Delta_{14})\rangle_y,\\ \langle D(\Delta_{21}, \Delta_{31}, \Delta_{41})\rangle_y &= \varepsilon_1\langle E(\Delta_{12}, \Delta_{13}, \Delta_{14})\rangle_x.\end{aligned} \quad (20)$$

In (19) and (20), $\Delta_{\vartheta 1} = -\Delta_{1\vartheta}$ ($\vartheta = 2, 3, 4$) according (4).

Symmetry relations (19) and (20) can be combined into one relationship that establishes a correlation between the average energy values:

$$\langle W(\Delta_{12}, \Delta_{13}, \Delta_{14})\rangle = \langle W(\Delta_{21}, \Delta_{31}, \Delta_{41})\rangle, \quad (21)$$

where $\langle W\rangle = \langle D\rangle_x\langle E\rangle_x + \langle D\rangle_y\langle E\rangle_y$.

Equality (21) can be treated (in terms of [11]) as the relation between the energies of two systems: initial and reciprocal to it. Note, in conclusion, that the symmetry relations are valid for piecewise homogeneous media with different structures. Their practical value is that they make it possible to verify and sometimes simplify calculations. For example, in our case, it suffices to calculate two of the four electric field components, $\langle D\rangle_x$ and $\langle E\rangle_x$ or $\langle D\rangle_y$ and $\langle E\rangle_y$. The other two can be found from (19) and (20).

EFFECTIVE PERMITTIVITY TENSOR

The component $\varepsilon_{\text{eff }xx}$ of the tensor $\hat{\varepsilon}_{\text{eff}}$ is found from relationship (7). In this relationship, the average electric fields are given by formulas (13) and (17). Thus, we have

$$\varepsilon_{\text{eff }xx} = \varepsilon_1\frac{1 - \alpha(\Delta_{12}s_1 + \Delta_{13}s_2 + \Delta_{14}s_3) + \Delta_{12}\Delta_{13}G(r_1, r_2) + \Delta_{12}\Delta_{14}G(r_1, r_3) + \Delta_{13}\Delta_{14}G(r_2, r_3) + \Delta_{12}^2H(r_1) + \Delta_{13}^2H(r_2) + \Delta_{14}^2H(r_3)}{1 + \beta(\Delta_{12}s_1 + \Delta_{13}s_2 + \Delta_{14}s_3) + \Delta_{12}\Delta_{13}U(r_1, r_2) + \Delta_{12}\Delta_{14}U(r_1, r_3) + \Delta_{13}\Delta_{14}U(r_2, r_3) + \Delta_{12}^2V(r_1) + \Delta_{13}^2V(r_2) + \Delta_{14}^2V(r_3)}, \quad (22)$$

where α and β are numerical parameters ($\alpha = 4/3$) such that $\alpha + \beta = 2$. Note that in composites that are isotropic on the average, $\alpha = \beta = 1$. For materials that become anisotropic after averaging, $\alpha \neq \beta$ but the strict equality $\alpha + \beta = 2$ holds. The numerical values of α and β apparently characterize the microscopic anisotropy of composites and are found during the averaging of the fields (it suffices to find one of them).

The parameters s_ν ($\nu = 1, 2, 3$) are the concentrations of the components in the medium. They are defined as the ratio of the cross-sectional area of the inclusions that falls into the averaging triangle to the area of this triangle (Fig. 2). We have

$$s_\nu^* = \frac{s_\nu/6}{s_{ABC}} = \frac{\pi r_\nu^2/6}{\sqrt{3}h^2/4} = \frac{2\pi}{3\sqrt{3}}r_\nu^{*2}, \quad (23)$$

where s_{ABC} is the area of the triangle ABC in Fig. 2 (later on, the asterisks will be omitted).

Note that the parameters α and β are related to the parameters p and q in (13) and (17) by the simple formulas $\alpha = 3\sqrt{3}\,p/2\pi$ and $\beta = 3\sqrt{3}\,q/2\pi$. The functions depending on the radii r_ν involved in (22) have been determined above. The other tensor component, $\varepsilon_{\text{eff }yy}$, is obtained from the second relationship in (7) and also from the reciprocity relation

$$\varepsilon_{\text{eff }xx}(\Delta_{12}, \Delta_{13}, \Delta_{14})\varepsilon_{\text{eff }yy}(\Delta_{21}, \Delta_{31}, \Delta_{41}) = \varepsilon_1^2, \quad (24)$$

which follows from symmetry relations (19) and (20).

Small Concentration of Inclusions

If the concentration of each of the inclusions is small and, hence, the radii of the cylindrical fibers are also small, the functions $U(\ldots)$, $G(\ldots)$, $V(\ldots)$, and $H(\ldots)$ are negligibly small, as follows from the calculations. Then, the terms containing the products and squares of the parameters $\Delta_{1\vartheta}$ ($\vartheta = 2, 3, 4$) in formula (22) can be neglected. As a result, the fractional rational expression for the component $\varepsilon_{\text{eff }xx}$ takes the fractional linear form

$$\varepsilon_{\text{eff }xx} = \varepsilon_1\frac{1 - \alpha(\Delta_{12}s_1 + \Delta_{13}s_2 + \Delta_{14}s_3)}{1 + \beta(\Delta_{12}s_1 + \Delta_{13}s_2 + \Delta_{14}s_3)}. \quad (25)$$

It is assumed that the parameters $\Delta_{1\vartheta}$ may take any arbitrary values. If we also assume that the permittivities of the matrix and the inclusions differ insignificantly (that is, the absolute values of the parameters $\Delta_{1\vartheta}$ are small), expression (25) can be linearized in these parameters:

$$\varepsilon_{\text{eff }xx} = \varepsilon_1[1 - 2(\Delta_{12}s_1 + \Delta_{13}s_2 + \Delta_{14}s_3)]. \quad (26)$$

Here, we took into account that $\alpha + \beta = 2$. Formula (26) defines the effective permittivity of a weakly inhomogeneous medium. It is accurate up to the terms of first order of smallness in s_ν and $\Delta_{1\vartheta}$. Such media no longer have the hexagonal structure, the composite as a whole becomes isotropic, and its permittivity becomes scalar: $\varepsilon_{\text{eff }xx} = \varepsilon_{\text{eff }yy} = \varepsilon_{\text{eff}}$.

Linear relationship (26) characterizes the effective permittivity of a matrix inhomogeneous material where the interaction between the inclusions can be ignored.

In this system, the local electric field inside the cylindrical inclusions is uniform, as follows from the associated equations.

If in a weakly inhomogeneous media, $\Delta_{12}s_1 + \Delta_{13}s_2 + \Delta_{14}s_3 = 0$, then, according to (26), the effective permittivity becomes equal to the permittivity of the matrix although the material contains the dissimilar inclusions. Physically, this means that in an inhomogeneous composite, the polarizations may cancel each other if the combination of the materials and concentrations of inclusions meets this condition. This is illustrated by the following example. Let the concentrations of the inclusions be the same: $s_1 = s_2 = s_3$. Then, $\Delta_{12} + \Delta_{13} + \Delta_{14} = 0$. The latter equality may take place if at least one of the parameters $\Delta_{1\vartheta}$ ($\vartheta = 2, 3, 4$) is negative and, hence, at least one of the inclusions has the permittivity exceeding that of the matrix. With several inclusions with permittivities larger and smaller than that of the matrix present, their polarizations have opposite signs and cancel if the above equalities are satisfied.

It is interesting that under these conditions, formula (25) for the effective permittivity gives the permittivity of the matrix. Only in this specific case the anisotropy of the medium in which the parameters $\Delta_{1\vartheta}$ are not necessarily small is absent.

Two-Component Dielectrics

The formula for the effective permittivity of two-component dielectrics is a special case of general expression (22). By varying the parameters $\Delta_{1\vartheta}$ ($\vartheta = 2, 3, 4$) and the radii of the inclusions r_ν ($\nu = 1, 2, 3$), one can derive several formulas for different structures. Of them, the case where a two-component dielectric retains the hexagonal configuration is the most interesting. Such a situation is realized when all the inclusions have the same radii, $r_1 = r_2 = r_3$, and permittivities, $\varepsilon_2 = \varepsilon_3 = \varepsilon_4$ ($\Delta_{12} = \Delta_{13} = \Delta_{14}$). Then, we have from (22)

$$\varepsilon_{\text{eff }xx} = \varepsilon_1 \frac{1 - 4\Delta_{12}s + \Delta_{12}^2 L(r)}{1 - 4\Delta_{12}s + \Delta_{12}^2 M(r)}. \qquad (27)$$

Here, $L(r) = 3[G(r) + H(r)]$ and $M(r) = 3[U(r) + V(r)]$, where $r = r_\nu$ and $s = s_\nu$. Such a dielectric has been studied in [6, 7] with the Rayleigh method. In those works, the following expression for the component $\varepsilon_{\text{eff }xx}$ of the tensor $\hat{\varepsilon}_{\text{eff}}$ was deduced (it is represented in terms of this work):

$$\varepsilon_{\text{eff }xx} = \varepsilon_1 \frac{1 - \Delta_{12}s_0 - \Delta_{12}^2 s_0^6[a + (b + c)s_0^6] + b\Delta_{12}^3 s_0^{13} + bc\Delta_{12}^4 s_0^{24}}{1 + \Delta_{12}s_0 - \Delta_{12}^2 s_0^6[a + (b + c)s_0^6] - b\Delta_{12}^3 s_0^{13} + bc\Delta_{12}^4 s_0^{24}}, \qquad (28)$$

where $s_0 = 3s$ is the total concentration of the inclusions and $a = 0.075422$, $b = 1.060283$, and $c = 0.000076$ are numerical coefficients. Solution (28) is obtained up to the terms of fourth order of smallness (in Δ_{12}).

The accuracy of the solutions obtained by the Rayleigh method has been thoroughly analyzed by Manteufel and Todreas [7]. They examined the analytical solutions accurate to from first to eleventh order of smallness and found, in particular, that the effective permittivities calculated by (28) and the numerical calculations differ by less than 0.1% for arbitrarily selected $|\Delta_{12}|$ if the concentration of the inclusions does not exceed $s_0 \approx 0.8$. For strongly inhomogeneous media, where the parameters s_0 and $|\Delta_{12}|$ are extremely high: $s_0 \longrightarrow s_{0\max}$ (where $s_{0\max}$ is the concentration when the inclusions are most close-packed, $s_{0\max} = \pi/2\sqrt{3}$) and $|\Delta_{12}| \longrightarrow 1$, the Rayleigh method requires higher order approximations.

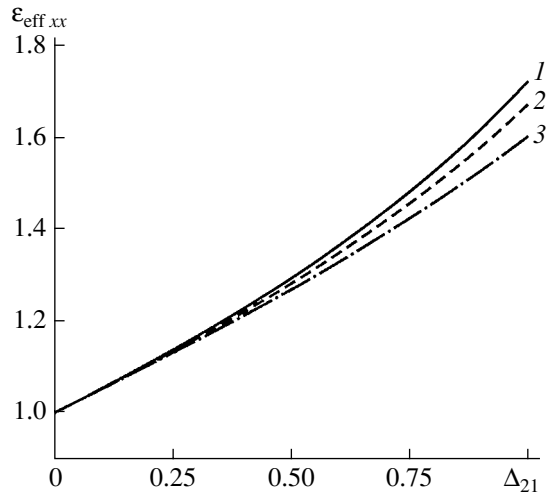


Fig. 3. $\varepsilon_{\text{eff }xx}$ vs. parameter Δ_{21} for the two-component fibrous dielectric. Calculation by formula (*1*) (27), (*2*) (28), and (*3*) 25. $\varepsilon_1 = 1$, $s_0 = 3s = 0.25$.

The accuracy of (28) is fairly good; therefore, comparing the results obtained with (28) and (27) under similar conditions, one can judge the accuracy of the single-dipole approximation used in evaluating the average parameters of dielectrics with the structure considered. To make the comparison more illustrative, we constructed the dependence $\varepsilon_{\text{eff }xx}(\Delta_{21})$ for $s_0 = 3s = 0.25$, 0.5, and 0.75 at $\varepsilon_1 = 1$ (Figs. 3–5). Here, the continuous curves refer to formula (27); dashed curves, to (28); and dot-and-dash lines, to (25). The last formula is the simplest and, hence, the most suitable for estimations. For $s_0 = 0.75$, formulas (27) and (28) are seen to yield virtually coinciding results throughout the range of the permittivity ε_2 (parameter Δ_{21}). At the smaller

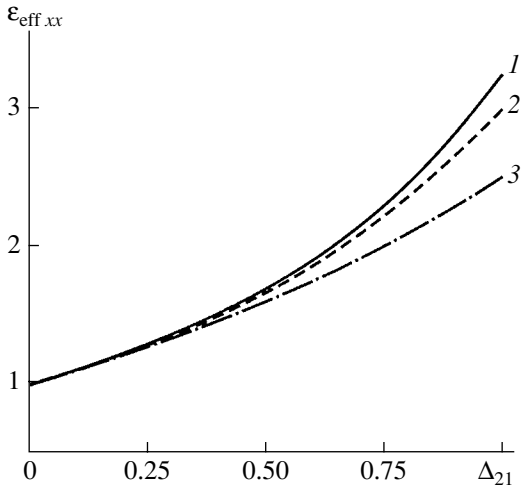


Fig. 4. The same as in Fig. 3 for $s_0 = 3s = 0.5$.

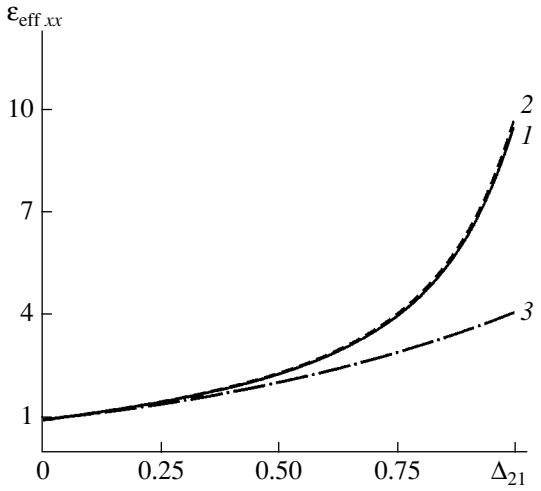


Fig. 5. The same as in Fig. 3 for $s_0 = 3s = 0.75$.

concentrations, formula (27) somewhat overestimates the value of $\varepsilon_{\text{eff}\,xx}$, starting from $\Delta_{21} = 0.5$ ($\varepsilon_2 \geq 3$). The error rises with increasing Δ_{21} and tends to the maximum in the limit $\varepsilon_2 \longrightarrow \infty$ ($\Delta_{21} = 1$). It follows from the plots that formula (25) is fairly accurate when the concentrations of the inclusions are small and their permittivities differ from that of the matrix only slightly.

CONCLUSION

The high order of symmetry, inherent in hexagonal packing, allowed us to analytically study a four-component dielectric material. The system has the regular structure and features the operation of mirror reflection from the x axis and the operation of discrete rotation (the triad axis): the system is brought into coincidence with itself under rotation through 120° (Fig. 1).

As to applications, four-component dielectrics offer the greater variability in physical properties than the two-component materials because of the great diversity of polarization phenomena in the former. As has been shown in this work, the effective permittivity of a four-component composite can be made equal to that of the matrix with the concentrations and permittivities of the inclusions properly selected. In two-component dielectrics, such properties are lacking.

APPENDIX A

Here, we give the formula for calculating the parameter q and the expressions for the functions $M(\ldots)$, $N(\ldots)$, and $V(\ldots)$ appearing in formulas (10)–(13). The parameter q is given by

$$q = 2\Bigg\{\sum_{m=1}^{\infty}\left(\frac{1}{3m-1}-\frac{1}{3m-2}\right)$$
$$+2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Bigg\{\frac{3m-1}{(3m-1)^2+3n^2}-\frac{3m-2}{(3m-2)^2+3n^2} \tag{A1}$$
$$+2\left[\frac{6m-5}{(6m-5)^2+3(2n-1)^2}-\frac{6m-1}{(6m-1)^2+3(2n-1)^2}\right]$$
$$+4\left[\frac{m}{12m^2+(2n-1)^2}-\frac{m-1}{12(m-1)^2+(2n-1)^2}\right]\Bigg\}\Bigg\}.$$

Note that formula (A1) yields the exact value of q and is valid for all the approximations used in the calculation of the electric field, including the single-dipole one. The function $M(r)$ ($r = r_1, r_2, r_3$) has the form

$$M(r) = \sum_{m=1}^{\infty}\left(\frac{1}{3m-2}-\frac{1}{3m-1}\right.$$
$$\left.+\frac{1}{r+3m-1}+\frac{1}{r-3m+2}\right)$$
$$+2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Bigg\{\frac{3m-2}{(3m-2)^2+3n^2}-\frac{3m-1}{(3m-1)^2+3n^2}$$
$$+\frac{r-3m+2}{(r-3m+2)^2+3n^2}+\frac{r+3m-1}{(r+3m-1)^2+3n^2} \tag{A2}$$
$$+2\left[\frac{6m-1}{(6m-1)^2+3(2n-1)^2}-\frac{6m-5}{(6m-5)^2+3(2n-1)^2}\right.$$
$$\left.+\frac{2r-6m+1}{(2r-6m+1)^2+3(2n-1)^2}\right.$$

$$-\frac{2r+6m-5}{(2r+6m-5)^2+3(2n-1)^2}\Bigg]\Bigg\}.$$

For the function $N(r)$, we have

$$N(r) = \sum_{m=1}^{\infty}\left(\frac{1}{3m-1}-\frac{1}{3m-2}+\frac{1}{r-3m+1}+\frac{1}{r+3m-2}\right)$$
$$+2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Bigg\{\frac{3m-1}{(3m-1)^2+3n^2}-\frac{3m-2}{(3m-2)^2+3n^2}$$
$$+\frac{r-3m+1}{(r-3m+1)^2+3n^2}+\frac{r+3m-2}{(r+3m-2)^2+3n^2} \quad \text{(A3)}$$
$$+2\Bigg[\frac{6m-5}{(6m-5)^2+3(2n-1)^2}-\frac{6m-1}{(6m-1)^2+3(2n-1)^2}$$
$$+\frac{2r+6m-1}{(2r+6m-1)^2+3(2n-1)^2}$$
$$+\frac{2r-6m+5}{(2r-6m+5)^2+3(2n-1)^2}\Bigg]\Bigg\}.$$

The function $V(r)$ is written as

$$V(r) = \frac{4}{3}r^2\Bigg\{r\sum_{m=1}^{\infty}\left(\frac{1}{r^2+3m^2}+\frac{1}{r^2-9m^2}\right)$$
$$+\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Bigg\{\frac{r-3m}{(r-3m)^2+3n^2}+\frac{r+3m}{(r+3m)^2+3n^2}$$
$$+2\Bigg[\frac{2r-6m-3}{(2r-6m+3)^2+3(2n-1)^2} \quad \text{(A4)}$$
$$+\frac{2r+6m-3}{(2r+6m-3)^2+3(2n-1)^2}\Bigg]\Bigg\}\Bigg\}.$$

APPENDIX B

In (14)–(16), the parameter p and the functions $L(r)$ and $H(r)$ have the following representation. The numerical parameter p is given by

$$p = 2\Bigg\{1+2\sum_{m=1}^{\infty}p\Bigg[\frac{1}{1+27m^2}+\frac{1}{1-9m^2}$$
$$+\frac{2}{1+27(2m-1)^2}\Bigg]+\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Bigg[\frac{2(1+3n)}{(1+3n)^2+27m^2}$$
$$+\frac{4(5-6n)}{(5-6n)^2+27(2m-1)^2}+\frac{4(1+6n)}{(1+6n)^2+27(2m-1)^2} \quad \text{(B1)}$$
$$+\frac{2-3n}{(2-3n)^2+27(m-1)^2}+\frac{3n-1}{(3n-1)^2+27(m-1)^2}$$
$$+\frac{2-3n}{(2-3n)^2+27m^2}+\frac{1-3n}{(1-3n)^2+27m^2}\Bigg]\Bigg\}.$$

In the calculations, two other formulas for p are used; however, all yield the same value $p = 1.612$.

The function $L(r)$, where $r = r_1, r_2, r_3$, has the form

$$L(r) = \frac{2}{\sqrt{3}}\Bigg\{r\sum_{m=1}^{\infty}\Bigg[\frac{1}{r^2+(3m-1)^2}+\frac{1}{r^2+(3m-2)^2}\Bigg]$$
$$+\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Bigg\{\frac{r+\sqrt{3}n}{(r+\sqrt{3}n)^2+(3m-1)^2}$$
$$+\frac{r-\sqrt{3}n}{(r-\sqrt{3}n)^2+(3m-1)^2}+\frac{r+\sqrt{3}n}{(r+\sqrt{3}n)^2+(3m-2)^2}$$
$$+\frac{r-\sqrt{3}n}{(r-\sqrt{3}n)^2+(3m-2)^2}$$
$$+2\Bigg[\frac{2r+\sqrt{3}(2n-1)}{[2r+\sqrt{3}(2n-1)]^2+(6m-1)^2} \quad \text{(B2)}$$
$$+\frac{2r-\sqrt{3}(2n-1)}{[2r-\sqrt{3}(2n-1)]^2+(6m-1)^2}$$
$$+\frac{2r+\sqrt{3}(2n-1)}{[2r+\sqrt{3}(2n-1)]^2+(6m-5)^2}$$
$$+\frac{2r-\sqrt{3}(2n-1)}{[2r-\sqrt{3}(2n-1)]^2+(6m-5)^2}\Bigg]\Bigg\}\Bigg\}.$$

The expression for the function $H(r)$ is as follows:

$$H(r) = \frac{4r^2}{3\sqrt{3}}\Bigg\{r\sum_{m=1}^{\infty}\left(\frac{1}{r^2+9m^2}+\frac{1}{r^2-3m^2}\right)$$
$$+\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\Bigg\{\frac{r+\sqrt{3}n}{(r+\sqrt{3}n)^2+9m^2}+\frac{r-\sqrt{3}n}{(r-\sqrt{3}n)^2+9m^2} \quad \text{(B3)}$$
$$+2\Bigg[\frac{2r+\sqrt{3}(2n-1)}{[2r+\sqrt{3}(2n-1)]^2+9(2m-1)^2}$$

$$+\frac{2r-\sqrt{3}(2n-1)}{[2r-\sqrt{3}(2n-1)]^2+9(2m-1)^2}\Bigg]\Bigg\}\Bigg\}.$$

The parameter p and the functions $L(r)$ and $H(r)$, as well as the parameter q and the functions $M(r)$, $N(r)$, and $V(r)$, have been derived by integrating the complex expressions for dipoles involved in formulas (3) for the electric field and in similar expressions.

REFERENCES

1. *Modern Crystallography*, Ed. by B. K. Vaĭnshtein, A. A. Chernov, and L. A. Shuvalov (Nauka, Moscow, 1981; Springer-Verlag, Berlin, 1988), Vol. 4.
2. L. P. Khoroshun, B. P. Maslov, and P. V. Leshchenko, *Forecasting of Effective Properties of Piezo-Active Composite Materials* (Naukova Dumka, Kiev, 1989).
3. L. M. Brekhovskikh, *Waves in Layered Media* (Nauka, Moscow, 1973; Academic, New York, 1980).
4. *Mechanics of Composite Materias*, Ed. by G. P. Sendeckyj (Academic, New York, 1974; Mir, Moscow, 1978).
5. A. G. Fokin, Usp. Fiz. Nauk **166**, 1069 (1996) [Phys. Usp. **39**, 1009 (1996)].
6. W. T. Perrins, D. R. McKenzie, and R. C. McPhedran, Proc. R. Soc. London, Ser. A **369**, 207 (1979).
7. R. D. Manteufel and N. E. Todreas, Int. J. Heat Mass Transf. **37**, 647 (1994).
8. Lord Rayleigh, Philos. Mag. **34**, 481 (1892).
9. Yu. P. Emets and Yu. P. Onofrichuk, IEEE Trans. Dielectr. Electr. Insul. **3**, 87 (1996).
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).
11. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **81**, 665 (1981) [Sov. Phys. JETP **54**, 355 (1981)].

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 58–63. Translated from Zhurnal Tekhnichesko ĭ Fiziki, Vol. 72, No. 1, 2002, pp. 60–65.

SOLID-STATE ELECTRONICS

Nanooxidation of *n*-$In_{0.53}Ga_{0.47}As$ with an Atomic Force Microscope

D. V. Sokolov

Institute of Analytical Instrument Making, Russian Academy of Sciences, Rizhskiĭ pr. 26, St. Petersburg, 198103 Russia

e-mail: dmitri sokolov@mail.ru

Received January 24, 2001

Abstract—It is shown that the features formed when an InGaAs layer lattice-matched with the InP substrate is modified by an atomic force microscope are of the oxide nature. Various models of InGaAs nanooxidation by an atomic force microscope are considered. The conventional Cabrera–Mott model refined by introducing the space charge into the oxide is taken as the basic model of nanooxidation by an atomic force microscope. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The knowledge of local oxidation mechanisms and kinetics is of crucial importance for optimizing the oxidation process by means of scanning probe microscopes, in particular, atomic force microscopes (AFMs). AFM oxidation is used in the production of electron devices with feature sizes as small as tens of nanometers. Also, an understanding of the local oxidation mechanism and kinetics would provide the reliable control of the device performance.

In many works on AFM oxidation (see, e.g., [1–3]), it has been reported that the geometric parameters of the resulting structures depend on the AFM probe size. This means that the electric field has an effect on the process. The high voltage increases the height and the width of the oxide lines because of enhanced anodizing or ionic diffusion.

A number of mechanisms are responsible for nanopatterning the substrate surface with an AFM. They depend on the polarity of the applied voltage and the process environment. For silicon, the AFM surface modification can be described within two models. One is based on the hypothesis that the oxidation near the probe–surface contact is facilitated by the electric field in the presence of oxygen [4, 5]. In this case, the probe is positively biased. The other model considers the anodizing of the surface covered by a thin water film under the action of the negatively biased probe [6]. Anodizing of semiconductors may proceed in either an oxygen plasma [7] or an electrolytic solution [8, 9].

At present, researchers suggest two mechanisms of local oxidation with the negatively biased probe that work in parallel. These are anodizing through a thin water film adsorbed on the surface [10] and electric-field-enhanced oxidation, where the electric field promotes the diffusion of ionized water molecules through the native oxide [3].

Attempts [11] to treat probe oxidation in terms of the Cabrera–Mott field model have failed [12], since the actual oxide growth rates much differ from the values predicted for anodizing. Yet, the Cabrera–Mott model shows the best correlation with the experiment and the discrepancy between the growth rates can be viewed as resulting from nonperfect oxidation conditions due to factors not taken into account. Such factors may be (1) mechanical stresses arising in the oxide because of a large difference in the oxide and silicon volumes [1], (2) space charge induced immediately at the point of potential application, or (3) a change in the field strength in response to a change in the shape and/or in the conductivity of the probe. These changes may be related to the plasticity of conductive coatings covering the probe and/or to the oxidation of the probe material itself.

Comparative studies of the effect of the probe polarity on the local oxidation of silicon [4] and chromium [13] have shown that the oxidized lines (features) are distinct, i.e., high, if the AFM probe is negatively biased. This has played a crucial role in deciding on the probe polarity for the formation of oxides more than 5 nm high. Such a height is necessary for the direct patterning of tunnel barriers in quantum devices.

The need for a water film in local oxidation has been demonstrated [14] in support of the fact that local oxidation has much in common with electrochemical anodizing. The optimization of the air humidity to produce oxides of a height of more than 5 nm has been reported [15].

An important parameter that characterizes the oxide growth rate is the norm of oxide growth, or the increment of the oxide thickness per volt. For a number of modifiable materials, the growth norm has been determined from experiments on anodizing at the macro-

1063-7842/02/4701-0058$22.00

level, so that it may indicate to what extent the nanoprocess and the macroprocess correlate.

EXPERIMENT

The subject of investigation was the *n*-$In_{0.53}Ga_{0.47}As$ layer ($n = 5 \times 10^{18}$ cm^{-3}) MBE-grown on the semi-insulating InP substrate. The surface conditioning procedure and the experimental setup used in the experiments were described in [15].

First, we tested the potentialities for oxidizing the *n*-InGaAs surface by applying a potential $V_{tip} = -5$ V to the tip when scanning a surface area of 100×100 nm in size. The scanning frequency and pitch were 4.3 Hz and 1 nm, respectively. After oxide dots and lines have been formed under various conditions, we studied the characteristics of AFM oxidation. The oxide dots were produced by applying a negative voltage from 1 to 10 V to the fixed tip for $t_{ox} = 60$ s. The lines were produced by scanning a 0.2- to 2.0-μm-long segment with a negative voltage $V_{tip} = 5$–30 V applied to the tip. The scanning rate was 10–260 nm/s. The identification of the structures obtained was performed immediately after the oxidation by imaging them with the same oxidizing tips. We also tested the feasibility of transferring the oxidized pattern into the InGaAs layer by etching the sample in an HF solution.

To reveal the basic laws of local AFM-assisted InGaAs oxidation, the same sample was totally oxidized (anodized) in a glycol–water solution (GWS) of tartaric acid. The anodizing solution and the associated setup were described in [16]. A transparent glass test tube was filled with the GWS, and a positive (relative to the Pt electrode) voltage of 15–30 V was applied to the sample. The sample was illuminated by an incandescent lamp to saturate the semiconductor surface by charge carriers.

The solution was prepared as follows. Assuming that the thin water film covering the sample is a neutral medium, a 3% water solution of tartaric acid ($C_4H_6O_6$) was neutralized by ammonium hydroxide to provide pH 6. The resulting solution was then mixed with propylene glycol so that the glycol-to-water volume ratio was 1 : 2.

RESULTS AND DISCUSSION

1. AFM Oxidation and Etching

Figure 1a shows the AFM image of the oxidized 100×100-nm surface area when the tip was under a negative voltage of 5 V. The cross section *A*–*A*' is depicted in Fig. 1b. The direction of the slow scanning was from left to right and from top down, as indicated by arrows in Fig. 1a. The size of the oxide island exceeds that of the scan area. The base of the island is 60 nm longer than the scan length. Moreover, the island is pentagonal (Fig. 1a) rather than rectangular, although the rectangular area was scanned. Finally, the height

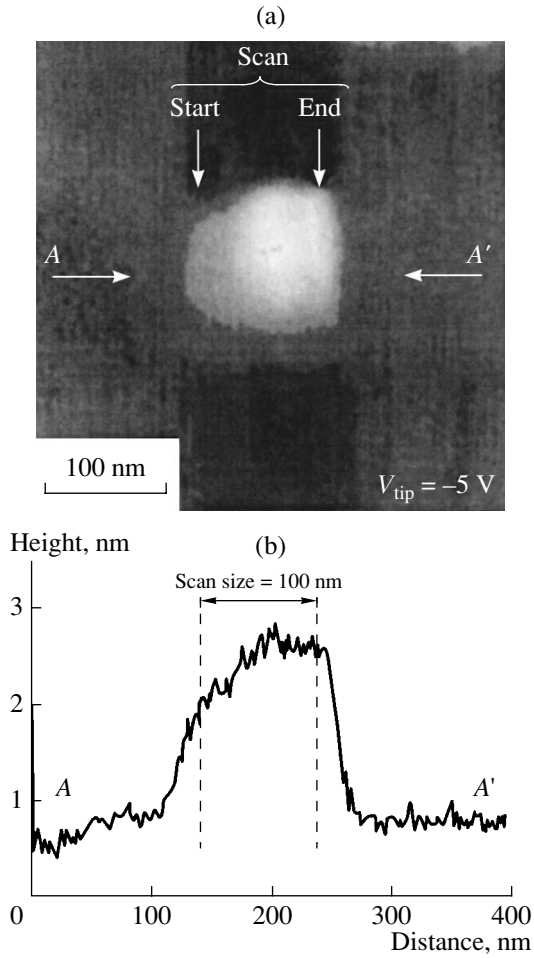


Fig. 1. (a) AFM image of the oxide island formed by scanning the 100×100-nm surface area and (b) the cross section of the island along the line A–A' of slow scanning.

(thickness) of the oxide smoothly varies over the former half of the scan area, remaining almost constant (about 1.8 nm) over the latter half.

In the contact mode of AFM operation (Fig. 2), the range of the threshold voltage (2–11 V) was close to that (5–10 V) observed for the AFM oxidation of GaAs—the material that is similar to InGaAs in composition [17, 18]. Such a spread in the threshold voltage can be related to one more factor essential for AFM oxidation: the condition of the tip during the oxidation. As follows from the experiments, the lowest threshold value, 2 V, corresponds to the early stage of scanning. After several tens of cycles including dot formation and imaging, the threshold voltage of oxidation rises to 10 V or higher. The stripping of the conductive layer from the AFM tip was corroborated by observing it in a scanning electron microscope (Fig. 3).

Figure 4a shows four lines oxidized by 15-fold scanning of the straight segment at the fixed frequency. In all the cases, the negative tip voltage was 5 V. The scan length was 0.2, 0.5, 1, and 2 μm from right to left. In

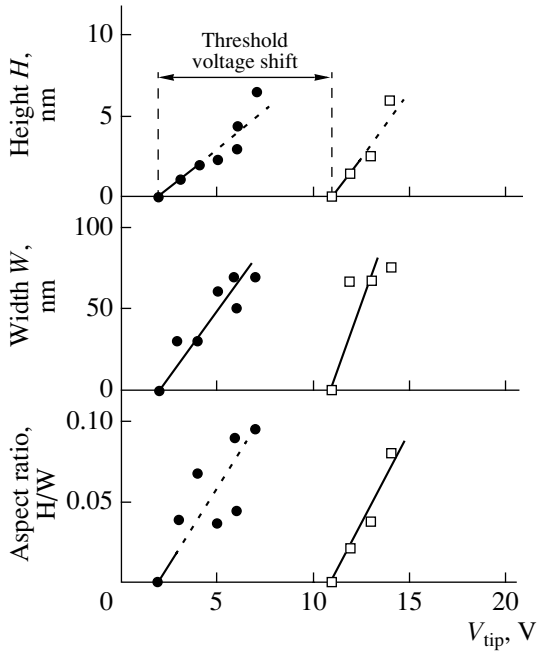


Fig. 2. Height, width, and aspect ratio vs. tip voltage for the dots oxidized in the contact mode.

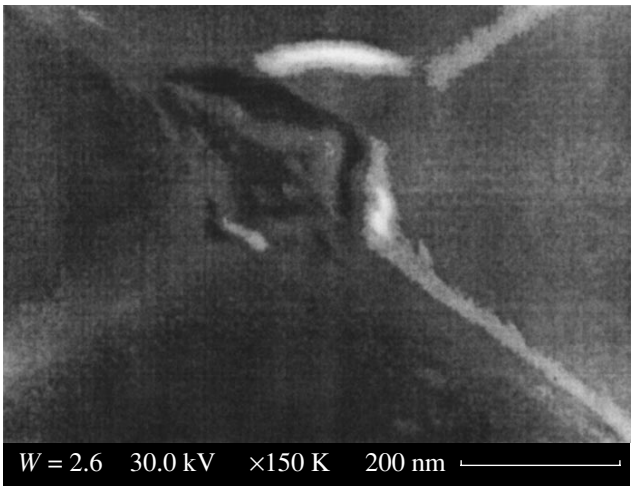


Fig. 3. SEM image of the end face of the Au-covered (30 nm) AFM tip after the operation in the contact mode. The conductive Au coating is stripped off at the intersection of the tip pyramid faces.

general, a decrease in the scan length at a fixed scanning frequency is equivalent to an increase in the time of scanning of an elementary surface area. The height and width of the oxide lines grow as the scan length decreases (Figs. 4b, 4c).

While the heights of the oxide dots markedly fluctuate, the geometrical sizes of the oxide lines remain virtually the same throughout the scan length. The slight change in the height of the lines is due to a recess on the base surface.

The fact that the height fluctuations for the lines are smaller than for the dots can be explained as follows. For a dot, the area of oxidation is small, so that the nonuniformity of the surface properties at the nanolevel becomes essential. Therefore, the necessary time of passivation varies from dot to dot. However, when the tip starts to scan a line, an oxidized area with a lateral size of several tens of nanometers forms at one of the scanning points by virtue of the tip geometry. As the tip passes this preoxidized area, further oxidation at each subsequent point of the tip path proceeds with a higher possibility. The easy anodizing of the preoxidized surface has also been observed in the case of GaAs [16].

Figure 4c plots the averaged height and width of each of the lines against the reciprocal scan length $1/L$. Since the scanning frequency and the number of passes of the tip were the same (0.13 Hz and 15, respectively) for all the lines, the reciprocal scan length $1/L$ gives the time of oxidation per unit length. As follows from Fig. 4c, the height quickly grows at small $1/L$ and then saturates near 4 nm at $1/L = 5$ μm^{-1}. It was found that such behavior can well be approximated by the following expression with a positive shift along the horizontal axis:

$$H = 4(\text{nm}) \times [1 - \exp\{-(1/L - 0.2)/0.8\}]. \quad (1)$$

The width of the oxide lines saturates more smoothly. The minimal width of the oxide, 65 nm, was observed at $L = 2$ μm.

The AFM oxides on the *n*-InGaAs surfaces were easily etched off by dipping into HF for 10 s. Figure 5a demonstrates the AFM images of the oxide line obtained at a negative tip voltage of 20 V and its associated groove. The cross sizes of both are shown in Fig. 5b. The oxide line is narrower than the groove, while the depth of the groove is 1.5 times larger than the height of the line. From the relationship between the height and the depth of the features, we found that the volumes of the initial InGaAs and the oxide formed relate as 3 to 5. As far as we know, literature data on an increase in the material volume upon InGaAs oxidation are lacking. However, our findings are totally or almost totally consistent with the data for volume increase upon the AFM oxidation of GaAs [17–19] and AlGaAs [20], as well as for macroscopic anodizing of GaAs [16]. The decrease in the linewidth after the oxidation points to the high selectivity of the etchant and is an indirect indication of the oxide nature of the features formed by AFM modification.

2. Oxidation Mechanisms

The majority of the results presented above can be explained in terms of the conventional theory of anodizing, which is based on the Cabrera–Mott model. According to this model, the thickness of an anodic oxide grows with time as [16]

$$x = x_\infty\{1 - \exp[-(t - t_p)/\tau]\}. \quad (2)$$

Here, x_∞ is the final oxide thickness in the steady state, t_p is the passivation time, and τ is the time constant of anodizing. The value of x_∞ is known to linearly increase with applied voltage and does not depend on the illumination intensity. The passivation time is the time of the so-called active-to-passive transition. This transition proceeds as follows: at the early stage of ionization, the anodizing products intensely (actively) dissolve. After a lapse of time t_p, the electrolyte becomes supersaturated by the reaction products in the immediate vicinity of the surface, which is the initial condition for passive oxidation.

The initially slow oxidation rate, which leads to the formation of the pentagonal oxide (Fig. 1b), as well as the positive shift along the horizontal axis in Fig. 3a, are due to the existence of the surface passivation time.

The linear dependence between the oxide dot height and the tip voltage (Fig. 2a), as well as the H–$1/L$ dependence for the oxide line (Fig. 4c), agree well with the model. Similar linear H–V dependences during AFM oxidation have also been observed by other authors. On the other hand, when the oxide lines were grown at different scanning rates v, different H–v relationships, such as $H \propto -\log(v)$, $1/H \propto \log(v)$, and $H \propto v^{-1/4}$, were found [11, 17, 21]. Bearing in mind that the scanning rate is proportional to the scan length in our experiments, we can represent the above relationships in the form similar to expression (1): $H \propto \log(1/L)$, $1/H \propto -\log(1/L)$, and $H \propto (1/L)^{1/4}$. In the works cited, the disagreement with the conventional Cabrera–Mott model was explained by the presence of mechanical stresses in the oxide. Our data (see Fig. 4), conversely, fit the model reasonably well and obey none of the above relationships.

Another disagreement with the conventional theory that was noted is the different growth norms for the oxide dots and lines. The maximal heights of the dots (Fig. 2) and lines (Fig. 4) were found to be 3 and 4.5 nm, respectively, at the fixed $V_{tip} = -5$ V. If is taken into account that the total thickness of the oxide is 2.5 times as large as the height of the features oxidized, the above values correspond to the growth norms 2.5 and 3.7 nm/V, respectively, for the dots and lines. According to the model, the growth norm must be the same in both cases, since it is limited only by electric-field-stimulated ionic diffusion into the oxide. Therefore, it is expected that there exists a mechanism reducing the electric field strength inside the dot. For such a mechanism, an increase in the H^+ ion concentration near the oxide layer during the oxidation has been considered [22]. When the oxide dots are formed, the positive voltage generates H^+ ions at the InGaAs–oxide interface according to the reaction

$$2In_{0.53}Ga_{0.47}As + 12h^+ + 6OH^- \longrightarrow 0.53In_2O_3 + 0.47Ga_2O_3 + As_2O_3 + 6H^+.$$

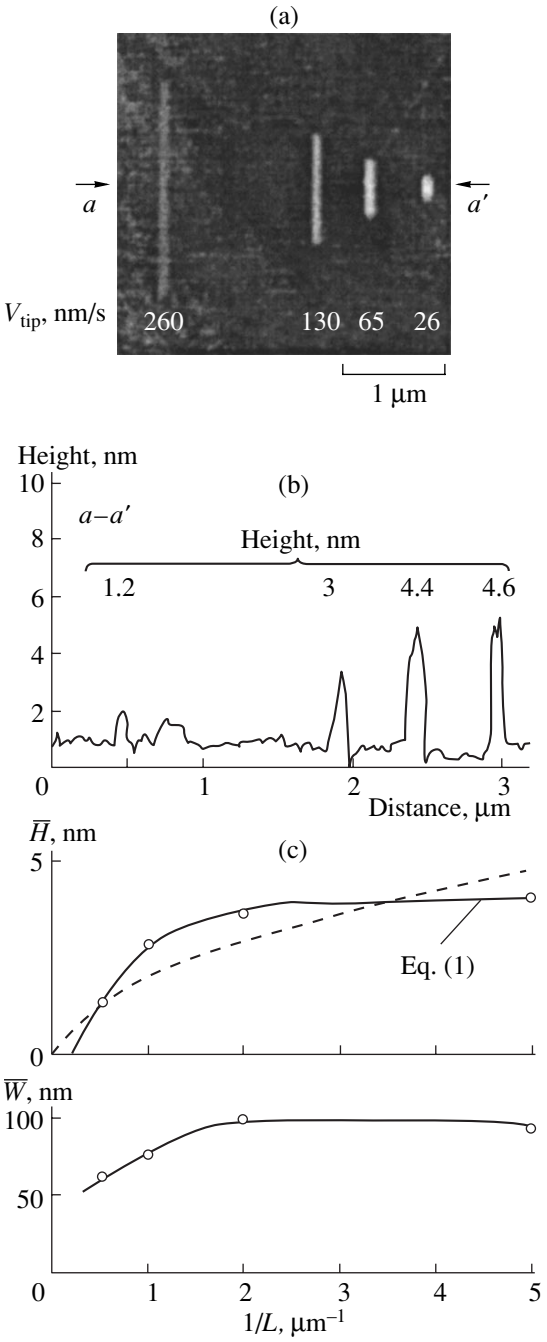


Fig. 4. (a) AFM images of the oxide lines formed by multiple scanning, (b) cross sections of the lines, and (c) heights and widths of the lines as a function of $1/L$.

After a lapse of time, the H^+ concentration in the oxide grows. A decrease in the electric field inside the oxide layer because of the screening effect of the H^+ ions may be a reason why the resulting oxidation norm is low.

It should be noted that we form the oxide lines by the 15-fold scanning of the surface with a frequency of 0.13 Hz. This amounts to the application of a pulse train

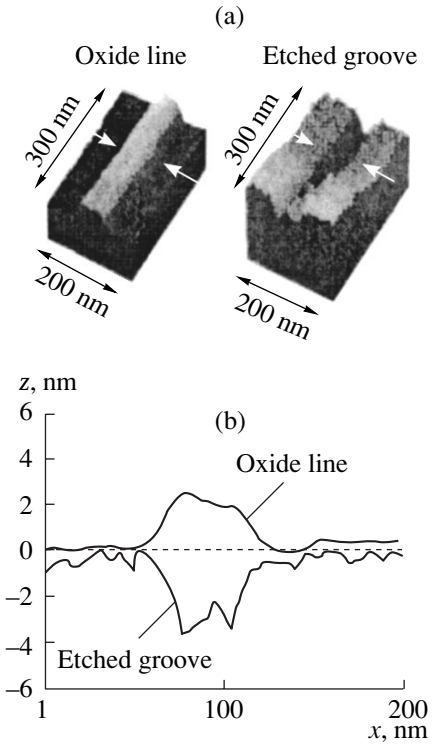


Fig. 5. (a) AFM image of the oxide line formed in the dark and its associated groove and (b) cross sections of the line and the groove.

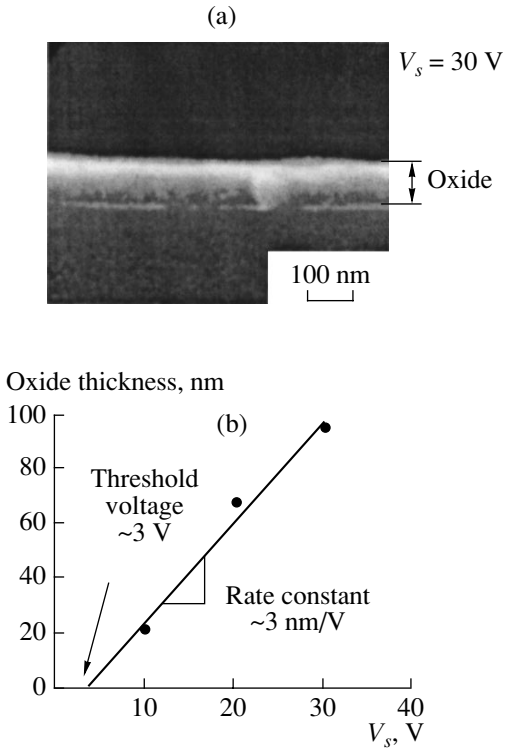


Fig. 6. (a) SEM image of the InGaAs layer anodized in the GWS and (b) InGaAs oxide thickness vs. InGaAs surface potential.

to each of the elementary segments of the line to be oxidized. In this situation, the H^+ ions generated at the instant a voltage pulse is applied to a given elementary segment diffuse from the oxide layer during the time when the pulse is off. The diffusion of the H^+ ions into the semiconductor decreases the ion concentration, thus providing high-rate long-term oxidation. Under these conditions, the oxide grows until its height reaches the value defined by the voltage applied. That is why the oxide lines are higher than the dots and the spread in the heights of the lines is smaller.

The even larger oxidation norm, about 3.9 nm/V, was observed when the oxidation was performed in the tapping mode (the probe vibrated with a frequency of about 180 kHz). Presumably, the vibrations of the probe having the constant positive potential relative to the surface caused the amplitude modulation of the electric field strength and, accordingly, led to the increase in the growth norm according to the mechanism of oxidation by multiple scanning. One may assume that the norm of InGaAs oxidation depends on the frequency of the ac voltage applied to the tip. For the dots, such a dependence was observed upon the AFM oxidation of Ti and Si [22, 23].

3. Macroanodizing in the GWS under Illumination

The thickness of the InGaAs layer macroanodized in the GWS was measured on the SEM image of its cross section. Figure 6a shows a typical SEM image of the cleaved surface when the anodizing voltage was 30 V and the initial current density, 1 mA/cm^3. From the dependence of the oxide thickness on the voltage applied (Fig. 6b), we find that the threshold voltage and the oxidation norm are about 2.5 V and 3 nm/V, respectively. The latter value approaches that for GaAs [16].

The closeness of the threshold voltages and the oxidation norms at macroanodizing and AFM oxidation in the contact mode with the tip worn out insignificantly is evidence for the similarity of the mechanisms behind both processes.

CONCLUSIONS

(1) Our findings on AFM nanooxidation are treated in terms of the Cabrera–Mott model refined (including the space charge induced during the oxidation).

(2) The shift of the threshold voltage for AFM oxidation toward larger values is associated with tip wear.

(3) The oxide growth norm increases with frequency of the tip voltage.

(4) The closeness of the threshold voltages and the oxidation norms at macroanodizing and AFM oxidation in the contact mode indicates the similarity of the mechanisms behind both processes.

REFERENCES

1. P. H. Avouris, T. Hertel, and R. Martel, Appl. Phys. Lett. **71**, 285 (1997).
2. P. J. Jorgensen, J. Chem. Phys. **37**, 874 (1962).
3. Y. Ejiri, M. Yasutake, and T. Hattori, in *Extended Abstracts of the International Conference on Solid State Devices and Materials, 1993*, p. 606.
4. J. A. Dagata, J. Schneir, H. H. Harary, *et al.*, Appl. Phys. Lett. **56**, 2001 (1990).
5. J. A. Dagata, J. Schneir, H. H. Harary, *et al.*, J. Vac. Sci. Technol. B **9**, 1384 (1991).
6. H. Sugimura, N. Kitamura, and H. Masuhara, Jpn. J. Appl. Phys., Part 2 **33**, L143 (1994).
7. J. R. Ligenza, J. Appl. Phys. **36**, 2703 (1965).
8. E. D. Palik, O. J. Glembocki, and J. Heard, Jr., J. Electrochem. Soc. **134**, 404 (1987).
9. O. J. Glembocki and R. E. Stelkbush, J. Electrochem. Soc. **132**, 145 (1985).
10. H. Suginura, T. Uchida, N. Kitamura, *et al.*, Jpn. J. Appl. Phys. **32**, L553 (1993).
11. T. Teuschler, K. Mahr, S. Miyazaki, *et al.*, Appl. Phys. Lett. **67**, 3144 (1995).
12. N. Cabrera and N. F. Mott, Rep. Prog. Phys. **12**, 163 (1949).
13. D. Wang, L. Tsau, K. L. Wang, *et al.*, Appl. Phys. Lett. **67**, 1295 (1995).
14. H. Sugimura, T. Uchida, N. Kitanura, *et al.*, J. Phys. Chem. **98**, 4352 (1994).
15. D. V. Sokolov, Nauchn. Priborostr. **11** (1), 15 (2001).
16. H. Hasegawa and H. L. Hartnagel, J. Electrochem. Soc. **123**, 713 (1976).
17. Y. Okada, S. Amano, M. Kawabe, *et al.*, J. Appl. Phys. **83**, 1844 (1998).
18. Y. Matsuzaki, K. Yuasa, J. Shirakashi, *et al.*, J. Cryst. Growth **201/202**, 656 (1999).
19. M. Ishii and K. Matsumoto, in *Extended Abstracts of the International Conference on Solid State Devices and Materials, 1995*, p. 953.
20. R. Held, T. Vancura, T. Heinzel, *et al.*, Appl. Phys. Lett. **73**, 262 (1998).
21. D. Stievenard, P. A. Fontaine, and E. Dubois, Appl. Phys. Lett. **70**, 3272 (1997).
22. J. A. Dagata, T. Inoue, J. Itoh, *et al.*, Appl. Phys. Lett. **73**, 271 (1998).
23. K. Matsumoto, Y. Gotoh, T. Maeda, *et al.*, Jpn. J. Appl. Phys. **38**, 477 (1999).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 6–12. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 8–14.

GASES AND LIQUIDS

Electromagnetic Radiation Produced by Linear Oscillations of a Charged Drop

A. I. Grigor'ev, S. O. Shiryaeva, A. S. Golovanov, and M. V. Rybakova

Yaroslavl State University, Sovietskaya ul. 14, Yaroslavl, 150000 Russia

e-mail: grig@uniyar.ac.ru

Received January 29, 2001

Abstract—A dispersion equation is derived for the capillary oscillations of a charged drop of an ideal incompressible liquid travelling in an ideal incompressible ambient allowing for the energy lost via emission of electromagnetic radiation. The intensity of electromagnetic radiation emitted by a single drop and a storm cloud has been estimated. *© 2002 MAIK "Nauka/Interperiodica".*

1. Investigations of capillary oscillations and stability of a charged drop are of considerable interest due to the diversity of scientific, technological, and industrial applications, in which the charged drop plays an important role (see, for example, [1–3] and the references therein). In particular, the issues associated with the emission of electromagnetic radiation by oscillating charged cloud droplets or raindrops are of interest in connection with the problems of radar sounding of meteorological objects [4–8]. The problem of calculating the intensity of electromagnetic radiation from an oscillating charged drop was originally formulated in [5], where the first estimates were also made for convective clouds. The importance of the problem calls for further elaboration of the theoretical concepts and, as well, for more precise estimates than those made in [5].

2. Consider a drop of an incompressible ideally conductive liquid with a density ρ_1 and surface tension σ carrying a charge Q surrounded by an ideal incompressible dielectric medium with a permitivity ε and density ρ_2. We shall seek the spectrum of capillary oscillations at the drop surface in the case where emission of electromagnetic radiation is possible as a result of displacements of the electric charge of the drop, under an assumption that vibrations of the drop surface do not change its center of mass position and volume.

An equation of the drop surface perturbed by a wave motion in a spherical coordinate system with the origin at the drop center of mass has the form

$$r(\Theta, t) = R + \xi(\Theta, t), \tag{1}$$

where R is the radius of the unperturbed drop and $\xi(\Theta, t)$ is the perturbation of the drop surface ($\max|\xi| \ll R$).

It will be assumed that the motion of fluids in the drop and the medium is of a potential type and the velocity fields in the drop and medium $\mathbf{U}_j(\mathbf{r}, t) = \nabla\varphi_j(\mathbf{r}, t)$ are fully defined by potential functions $\varphi_j(\mathbf{r}, t)$ (index $j = 1$ denotes quantities relating to the drop, and $j = 2$ to the external medium). Mathematical formulation of this problem in question has the form

$$\Delta\phi_j(\mathbf{r}, t) = 0; \quad j = 1, 2; \tag{2}$$

$$\Delta\mathbf{E}(\mathbf{r}, t) = \frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}; \quad \mathrm{div}\mathbf{E} = 0; \tag{3}$$

$$r \longrightarrow 0: \quad \phi_1(\mathbf{r}, t) = \mathrm{const}; \tag{4}$$

$$r \longrightarrow \infty: \quad \phi_2(\mathbf{r}, t) \longrightarrow 0; \tag{5}$$

$$\mathbf{E}(\mathbf{r}, t) \longrightarrow 0; \tag{6}$$

$$r = R + \xi(\Theta, t): \quad \frac{\partial\xi}{\partial t} = \frac{\partial\phi}{\partial n}; \tag{7}$$

$$\frac{\partial\phi_1}{\partial n_1} = \frac{\partial\phi_2}{\partial n_2} = \frac{\partial\phi}{\partial n}; \tag{8}$$

$$\Delta p - \rho_1\frac{\partial\phi_1}{\partial t} + \rho_2\frac{\partial\phi_2}{\partial t} + \frac{\varepsilon}{8\pi}(\mathbf{E})^2 = \sigma\left[\frac{2}{R} - \frac{1}{R^2}(2 + \hat{L})\xi\right]; \tag{9}$$

$$\boldsymbol{\tau}(\mathbf{r}, t)\mathbf{E}(\mathbf{r}, t) = 0. \tag{10}$$

In the relationships above $\mathbf{n}(\mathbf{r}, t)$ is a unit vector along the normal to the interface between the fluids; $\mathbf{n} \parallel \mathbf{n}_1$; $\mathbf{n}_1$ and $\mathbf{n}_2$ are unit vectors of the normal to the interface directed outward relative to the drop and the medium, respectively; $\boldsymbol{\tau}(\mathbf{r}, t)$ is a unit vector of a tangent to the drop surface; $\hat{L}$ is an angular component of the Laplacian in spherical coordinates; and c is the velocity of light.

To make the above set of equations complete, we add conditions of invariability of the drop volume and its center of mass position, as well as a conservation

condition for the total charge of the drop

$$\int_V r^2 dr \sin\Theta d\Theta d\phi = \frac{4}{3}\pi R^3,$$

$$V = [0 \le r \le R + \xi(\Theta, t),\ 0 \le \Theta \le \pi,\ 0 \le \phi \le 2\pi];$$

$$\oint_S \xi(\Theta, t)\mathbf{n}_r \sin\Theta d\Theta d\phi = 0; \qquad (11)$$

$$S = [r = R + \xi(\Theta, t),\ 0 \le \Theta \le \pi,\ 0 \le \phi \le 2\pi];$$

$$-\frac{1}{4\pi}\oint_S (\mathbf{n}\nabla\Phi) dS = \frac{Q}{\varepsilon}.$$

Below, all derivatives in the boundary conditions defined in a linear approximation with respect to $|\xi|/R$ refer to the unperturbed drop surface ($r = R$), as adopted in the theory of waves of infinitely small amplitude.

3. Let us represent the surface perturbation $\xi(\Theta, t)$ by a capillary wave of an initially equilibrium spherical drop as a series in Legendre polynomials $\Im_n(\cos\Theta) \equiv \Im_n(\mu)$ normalized as follows

$$\xi(\Theta, t) = \sum_{n=2}^{\infty} \alpha_n \Im_n(\mu) \exp(i\omega t).$$

We obtain

$$\int_1^1 \Im_n(\mu)\Im_m(\mu) d\mu = \delta_{nm}.$$

In the expression for $\xi(\Theta, t)$, the lower summation index is taken to be equal to 2 to meet the condition of invariability of the drop volume and center of mass position by excluding oscillations with $n = 0$ and $n = 1$, and ω is a complex frequency of capillary oscillations of the surface.

Analogously, in the form of a series in polynomials $\Im_n(\mu)$, we will be seeking solutions of Eq. (2) for the hydrodynamic potentials satisfying the boundary conditions (4) and (5)

$$\phi_1(r, \Theta, t) = \sum_{n=2}^{\infty} A_n r^n \Im_n(\mu) \exp(i\omega t), \qquad (12)$$

$$\phi_2(r, \Theta, t) = \sum_{n=2}^{\infty} B_n \frac{1}{r^{n+1}} \Im_n(\mu) \exp(i\omega t). \qquad (13)$$

The coefficients A_n and B_n are of the same order of smallness as α_n. The relationship between these coefficients and α_n is easily found from (7) and (8) as follows

$$A_n = -\alpha_n \frac{i\omega}{n} R^{1-n}; \quad B_n = \alpha_n \frac{i\omega}{(n+1)} R^{n+2}. \qquad (14)$$

To derive the dispersion equation for this problem, we differentiate the dynamic boundary condition (9) with respect to time (at Θ = const); then, taking into account (7) and (8), we get

$$-\rho_1 \frac{\partial^2\phi_1}{\partial t^2} + \rho_2 \frac{\partial^2\phi_2}{\partial t^2} + \sigma \frac{1}{R^2}(2 + \hat{L}) \frac{\partial\phi_1}{\partial r} + \frac{\partial}{\partial t}\left(\frac{\varepsilon}{8\pi}(\mathbf{E})^2\right) = 0. \qquad (15)$$

In Eq. (15) there remains to be determined the time derivative of the pressure exerted by the electrical field on the drop surface. To find it, it is necessary to solve an electrodynamic problem defined by Eqs. (3), (6), (10), and (11); this is done in the Appendix, where it is found that

$$\frac{\partial}{\partial t}\left(\frac{\varepsilon}{8\pi}(\mathbf{E})^2\right) = \frac{i\omega}{4\pi}\frac{Q^2}{\varepsilon R^5} \sum_n \alpha_n \Im_n(\mu) \exp(i\omega t) \times \left(n - 1 - \frac{i\omega R}{c}\frac{(n+1)}{n}\right). \qquad (16)$$

Now we substitute (12)–(14) and (16) into (15) and after simple transformations obtain the sought for dispersion equation

$$\omega_n^2 - \frac{i\omega_n Q^2 (n+1)^2}{[\rho_2 n + (n+1)\rho_1] 4\pi c\varepsilon R^5} - \frac{(n+1)n(n-1)}{[\rho_2 n + (n+1)\rho_1]} \times \frac{\sigma}{R^3}\left[(n+2) - \frac{Q^2}{4\pi\varepsilon R^3\sigma}\right] = 0,$$

which is quadratic in ω_n.

The coefficient of ω_n in the second term of this equation is equal to twice the damping factor of capillary waves in the considered system of ideal fluids

$$\eta \equiv \frac{Q^2 (n+1)^2}{2[\rho_2 n + (n+1)\rho_1] 4\pi c\varepsilon R^5}.$$

Thus, it is seen that the second term in the dispersion equation describes the energy lost by the capillary oscillations of the charged drop through radiation of electromagnetic waves. Note that the dimensionless combination of physical parameters

$$W \equiv \frac{Q^2}{4\pi\varepsilon R^3\sigma}$$

characterizes the drop stability with respect to its own charge [1]. According to Rayleigh's theory, the nth mode of the capillary oscillations of a drop becomes unstable, if the Rayleigh's criterion $W \ge (n + 2)$ is satisfied.

4. An expression for the radiation power emitted at a frequency ω_n can be written in the form [5]

$$I = -\frac{dw_n}{dt} \equiv w_n \frac{0.5Q^2(n+1)^2}{[\rho_2 n + (n+1)\rho_1]4\pi c\varepsilon R^5}, \qquad (17)$$

where w_n is the energy of surface oscillations of the nth mode. This energy, according to the virial theorem, is equal to twice the kinetic energy of the internal motion of the liquid inside the drop averaged over the oscillation period; for the nth mode of oscillations this energy is [9]

$$w_n = \frac{2\pi R^3 \rho_1 \omega_{n0}^2 \alpha_n^2}{n(2n+1)},$$
$$\omega_{n0}^2 = \frac{(n+1)n(n-1)}{[\rho_2 n + (n+1)\rho_1]}\frac{\sigma}{R^3}[(n+2) - W]. \qquad (18)$$

Now we substitute (18) into (17) and write the final expression for the intensity of the electromagnetic radiation from a single vibrating charged drop of an ideal liquid in the form

$$I = \frac{(n+1)^2\rho_1}{n(2n+1)[\rho_2 n + (n+1)\rho_1]}\frac{Q^2\omega_{n0}^2\alpha_n^2}{4c\varepsilon R^2}$$
$$\equiv \frac{(n+1)^2(n^2-1)\rho_1}{(2n+1)[\rho_2 n + (n+1)\rho_1]^2}\left(\frac{Q^2}{R^3}\right)\left(\frac{\alpha_n^2}{R^2}\right) \qquad (19)$$
$$\times \frac{\sigma}{4c\varepsilon}[(n+2) - W].$$

Let us consider two most probable sources of electromagnetic radiation from droplets in a storm cloud.

(a) The first probable radiation source is the vibrations of small charged droplets of sizes from 3 to 30 μm encountered in a cloud most frequently. The concentration of such droplets in a cloud is ~10^3 cm^{-3}. Vibrations of cloud droplets can be caused by various factors, such as aggregation, breaking up into small droplets as a result of collisions or electrostatic instability, hydrodynamic or electrical interaction with droplets passing close by, aerodynamic interaction with developed small-scale turbulences typical for storm clouds. The oscillation amplitudes α_n of cloud droplets, according to observations under natural conditions [6–8], can reach a few tens of percent of the droplet radius. In other words, the ratio (α_n/R) involved in (19) can be assumed independent of the droplet radius and equal to a few tenths of unity. Data on the average charges carried by droplets in storm clouds measured under natural conditions [10] show that for droplets of different sizes the ratio (Q^2/R^3) can roughly be assumed constant. The magnitude of parameter W for an average droplet charge, according to [10], is much less than unity; that is, most of the droplets are rather far from the instability threshold related to their charge. On the whole, at $\rho_1 \gg \rho_2$, according to (19), the intensity of electromagnetic radiation emitted by a vibrating cloud droplet depends first of all on the mode number n and the relative amplitude (α_n/R) of vibrations. For numerical evaluations we take $\varepsilon = 1$, $n = 2$, $\sigma = 73$ dyn/cm, $\rho_1 = 1$ g/cm^3, $\rho_2 = 1.3 \times 10^{-3}$ g/cm^3, $\alpha_n = 0.1R$, $R = 30$ μm, $Q = 2.5 \times 10^{-6}$ CGSE. It is then easy to estimate from (19) that $I \approx 10^{-21}$ W and the radiation frequency is ≈100 kHz. The intensity of electromagnetic radiation from a volume of 1 cm^3 will be roughly a thousand times as high, or ≈10^{-18} W, if one takes into account the abovementioned constancy of the ratio (Q^2/R^3), with the radiation extending over a frequency band from 100 kHz to 3 MHz. As for the total intensity of electromagnetic radiation from a storm cloud 5 km in diameter, its magnitude will be quite significant; namely, $I_{in} \approx 70$ mW. In our evaluation we assumed that the vibrations of the droplets occurred predominantly in the principal mode, $n = 2$. If one takes into account excitation of higher modes, then the total intensity will be higher in proportion to n^2.

(b) The second possible source of electromagnetic radiation, according to [5], is associated with collisions of the hydrometeors falling free through the cloud with finer droplets, causing continuous vibrations and, consequently, emission of electromagnetic radiation. However, in [5] it was proposed that the emitting hydrometeors could be highly charged drops with a radius of $R = 1$ mm whose concentration in clouds according to observations [10, 11] is very low, around 1 m^{-3}. Therefore, estimates of the electromagnetic radiation from a storm cloud based on this mechanism carried out for extremal magnitudes of the charges and concentrations of drops with $R = 1$ mm, appear to be much too high. Nevertheless, the mechanism proposed in [5] still proves valid if account is taken of the droplets an order of magnitude smaller in size, $R = 100$ μm, whose concentration in storm clouds is sufficiently high, about 10^3 m^{-3}, [10, 11], and the free fall velocity is ≈78 cm/s. Falling at this velocity, through a cloud of droplets with the radii from 3 to 30 μm, most within the range 3–7 μm, a hydrometeor will be experiencing about 22 collisions per second exciting in it oscillations of modes with $n \approx 20$. Taking $\alpha_n = 0.1R$, $Q = 5 \times 10^{-4}$ CGSE [10], it can be estimated by (19) that the intensity of electromagnetic radiation from a single hydrometeor is $I \approx 1.5 \times 10^{-16}$ W and its frequency is ≈25 kHz. The total intensity of electromagnetic radiation of all hydrometeors in a storm cloud 5 km in diameter is found to be equal to $I_{in} \approx 10$ mW.

Thus, of the two possible sources of electromagnetic radiation from droplets in a storm cloud discussed above the first, associated with low modes of oscillations in fine droplets, produces more intense radiation and at higher frequencies than the second one associated with oscillations of higher modes in large drops falling freely through a cloud in the gravitational field (hydrometeors).

It should be noted that all obtained estimates refer to an ideal conducting liquid and, therefore, are appreciably overvalued (at least, in the case of water). To obtain more realistic estimates it is necessary to take into account finite conductivity of real liquids or, which is the same, the relaxation of their charge. This, in turn, implies solving the problem for a viscous liquid, because only in this case it is possible to compensate tangential stresses at the interface arising due to the charge relaxation. This, however, is a subject for further studies.

CONCLUSION

Background radiation from charged clouds of natural or artificial origin can be partially explained by electromagnetic radiation produced as a result of capillary oscillations of the droplets in the cloud, both large, $R \approx 100$ μm, and small, $R \approx 3$ μm, using a linear approximation with respect to the amplitudes of the oscillation modes.

APPENDIX

Calculations of the Electrical Field Strength near the Surface of a Vibrating Charged Drop and the Time Derivative of the Pressure Exerted by the Electrical Field on the Drop Surface

1. To describe the electrical field around a charged conductive spherical drop whose surface is being perturbed by capillary waves, it is necessary to solve a set of equations involving a continuity condition for the electrical field and the wave equations

$$\mathrm{div}\mathbf{E} = 0, \quad \Delta\mathbf{E} = \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t^2} \tag{1A}$$

with the boundary conditions

$$r \longrightarrow \infty: \quad \mathbf{E}(\mathbf{r}, t) \longrightarrow 0;$$

$$r = R + \xi: \quad \boldsymbol{\tau}(\mathbf{r}, t)\mathbf{E}(\mathbf{r}, t) = 0,$$

$$-\frac{1}{4\pi}\oint_S (\mathbf{n}\nabla\Phi)dS = \frac{Q}{\varepsilon},$$

$$S = [r = R + \xi(\Theta, t),\ 0 \le \Theta \le \pi,\ 0 \le \phi \le 2\pi].$$

We shall seek the solution of Eqs. (1A) in the form of a superposition in spherical coordinates with the origin at the center of mass of the drop

$$\mathbf{E} = \mathbf{E}_0(r) + \mathbf{E}_1(r, \Theta, t), \tag{2A}$$

where $\mathbf{E}_0(r)$ is the electrical field strength in the vicinity of the unperturbed sphere and $\mathbf{E}_1(r, \Theta, t)$ is an additional contribution to the field strength from capillary oscillations at the drop surface; this contribution is of the same order of smallness as the surface perturbation $\mathbf{E}_1 \sim \xi$.

Due to linearity of Eqs. (1A) and the expansion (2A), the vectors $\mathbf{E}_0(r)$ and $\mathbf{E}_1$ will be solutions of sets of equations analogous to (1A).

2. To find the field strength in a zero approximation $\mathbf{E}_0(r)$ we should solve the problem

$$\mathrm{curl}\mathbf{E}_0 = 0; \tag{3A}$$

$$\mathrm{div}\mathbf{E}_0 = 0; \tag{4A}$$

$$r = R: \quad \mathbf{E}_0\boldsymbol{\tau} = 0; \tag{5A}$$

$$\oint_S \mathbf{E}_0 d\mathbf{S} = \frac{4\pi Q}{\varepsilon}; \tag{6A}$$

$$r \longrightarrow \infty: \quad \mathbf{E}_0(\mathbf{r}, t) \longrightarrow 0.$$

We shall seek $\mathbf{E}_0$ in the form $\mathbf{E}_0 = \nabla\Psi$. In this case, Eq. (3A) becomes identical and Eq. (4A) transforms into Laplace's equation for the potential Ψ

$$\Delta\Psi = 0. \tag{7A}$$

Due to the central symmetry of the problem in the zero approximation, Eq. (7A) takes the form of an ordinary differential equation of the second order

$$\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial\Psi}{\partial r} = 0,$$

with a solution

$$\Psi = -\frac{A}{r}.$$

Therefore, for the field strength $\mathbf{E}_0$ we obtain

$$\mathbf{E}_0 = \frac{A}{r^2}\mathbf{e}_r. \tag{8A}$$

The boundary condition (5A) is satisfied for arbitrary values of the constant A due to $\mathbf{e}_r \perp \boldsymbol{\tau}$. To find this constant let us substitute (8A) into the boundary condition (6A); then we obtain $A = Q/\varepsilon$. Finally, the expression for the field strength $\mathbf{E}_0$ takes the form

$$\mathbf{E}_0(r) = \frac{Q}{\varepsilon r^2}\mathbf{e}_r.$$

3. To determine the additional contribution of the first order of smallness to the electrical field strength we represent Eqs. (1A) in a scalar form. Let us represent vector $\mathbf{E}_1$ in the form of an expansion

$$\mathbf{E}_1 = \hat{\mathbf{N}}_1\Psi_1 + \hat{\mathbf{N}}_2\Psi_2 + \hat{\mathbf{N}}_3\Psi_3, \tag{9A}$$

where

$$\left.\begin{aligned} &\hat{\mathbf{N}}_1 = \nabla;\ \hat{\mathbf{N}}_2 = \nabla\times\mathbf{r};\ \hat{\mathbf{N}}_3 = \nabla\times(\nabla\times\mathbf{r}) \\ &\hat{\mathbf{N}}_1^+ = -\nabla;\ \hat{\mathbf{N}}_2^+ = \mathbf{r}\times\nabla;\ \hat{\mathbf{N}}_3^+ = (\mathbf{r}\times\nabla)\times\nabla \\ &\times\hat{\mathbf{N}}_j^+\hat{\mathbf{N}}_m = 0 \quad (m \ne j), \end{aligned}\right\} \tag{10A}$$

and Ψ_j is an unspecified scalar function.

Taking into account the property expressed by (10A), the first equation in the set (1A) transforms into Laplace's equation for the function Ψ_1

$$\begin{aligned}\nabla\mathbf{E}_1 &= -\hat{\mathbf{N}}_1^+\mathbf{E}_1 = -\hat{\mathbf{N}}_1^+(\hat{\mathbf{N}}_1\Psi_1 + \hat{\mathbf{N}}_2\Psi_2 + \hat{\mathbf{N}}_3\Psi_3)\\ &= -\hat{\mathbf{N}}_1^+\hat{\mathbf{N}}_1\Psi_1 = \nabla\nabla\Psi_1 = \Delta\Psi_1 = 0.\end{aligned} \tag{11A}$$

It is easy to see that the operators $\hat{\mathbf{N}}_j$ are commutative with Laplace's operator; that is, $\hat{\mathbf{N}}_j\Delta = \Delta\hat{\mathbf{N}}_j$, and therefore the second equation of the set (1A) transforms into a set of three scalar equations. Let us substitute the expansion (9A) into the wave equation of the set (1A)

$$\sum_{j=1}^{3}\hat{N}_j\left[\Delta\Psi_j - \frac{1}{c^2}\frac{\partial^2\Psi_j}{\partial t^2}\right] = 0.$$

Multiplying successively the last expression from the left by $\hat{\mathbf{N}}_j^+$ and taking into account that $\hat{\mathbf{N}}_j^+\hat{\mathbf{N}}_j \neq 0$, we obtain

$$\Delta\Psi_j - \frac{1}{c^2}\frac{\partial^2\Psi_j}{\partial t^2} = 0 \quad (j = 1, 2, 3). \tag{12A}$$

In solving a problem of electromagnetic radiation from a drop, it is natural to take $\Psi \sim \exp(-i\omega t)$; this reduces Eqs. (12A) to Helmholtz's equations.

From Eq. (12A) at $j = 1$ and Eq. (11A) we obtain $(\omega^2/c^2)\Psi_1 = 0$ and because the oscillation frequency ω is different from zero the scalar function $\Psi_1 = 0$. Thus, the electrical field $\mathbf{E}_1$ produced by the surface oscillations of a charged drop will be

$$\mathbf{E}_1 = \hat{\mathbf{N}}_2\Psi_2 + \hat{\mathbf{N}}_3\Psi_3. \tag{13A}$$

The functions Ψ_2 and Ψ_3 are solutions of a scalar Helmholtz's equation having the form

$$\Delta\Psi_j + \frac{\omega^2}{c^2}\Psi_j = 0; \quad j = 2, 3. \tag{14A}$$

In the general case, components of $\mathbf{E}_1$ in a spherical coordinate system, according to (13A) are written as

$$\hat{\mathbf{N}}_2\Psi_2 = 0\mathbf{e}_r + \frac{1}{\sin\Theta}\frac{\partial\Psi_2}{\partial\phi}\mathbf{e}_\Theta - \frac{\partial\Psi_2}{\partial\Theta}\mathbf{e}_\phi, \tag{15A}$$

$$\begin{aligned}\hat{\mathbf{N}}_3\Psi_3 &= -\frac{1}{r}\hat{L}\Psi_3\mathbf{e}_r + \frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial\Psi_3}{\partial\Theta}\mathbf{e}_\Theta\\ &+ \frac{1}{r\sin\Theta}\frac{\partial}{\partial r}r\frac{\partial\Psi_3}{\partial\phi}\mathbf{e}_\phi.\end{aligned} \tag{16A}$$

Because of the axial symmetry of the problem, the terms in expressions (15A) and (16A) involving derivatives with respect to ϕ cancel. Then we have

$$\begin{aligned}\hat{\mathbf{N}}_2\Psi_2 &\equiv -\frac{\partial\Psi_2}{\partial\Theta}\mathbf{e}_\phi,\\ \hat{\mathbf{N}}_3\Psi_3 &= -\frac{1}{r}\hat{L}\Psi_3\mathbf{e}_r + \frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial\Psi_3}{\partial\Theta}\mathbf{e}_\Theta.\end{aligned} \tag{17A}$$

4. A solution for the functions Ψ_j is written in the form

$$\Psi_j = \sum_n D_n^{(j)}(kr)^{-1/2}H_{n+1/2}^{(2)}(kr)\Im_n(\mu); \quad j = 2, 3, \tag{18A}$$

where $H_{n+1/2}^{(2)}(kr)$ is the second Hankel's function chosen in such a way that the electromagnetic wave would be asymptotically diverging at $r \longrightarrow \infty$; $k = \omega/c$ is the wave number.

The unknown constants $D_n^{(j)}$ in solutions (18A) are determined from the condition of equipotentiality of the drop surface and conservation of the total charge of the drop. Note that for the toroidal component of the field $\mathbf{E}_1$ defined by the function Ψ_2 these boundary conditions lead to the following relations

$$r = R + \xi: \quad (\hat{\mathbf{N}}_2\Psi_2)\boldsymbol{\tau} = 0, \tag{19A}$$

$$\oint_S[(\hat{\mathbf{N}}_2\Psi_2)\mathbf{n}]ds = 0, \tag{20A}$$

where $\boldsymbol{\tau}$ and $\mathbf{n}$ are the tangential and normal unit vectors of the drop surface, respectively.

As $\mathbf{E}_1$ and, according to (11A), the functions Ψ_j are of the first order of smallness, the expressions (19A) and (20A) should refer to the unperturbed drop surface, $r = R$. The vector $\boldsymbol{\tau}$ will then be defined by the unit vectors of the spherical coordinate system $\mathbf{e}_\Theta$ and $\mathbf{e}_\phi$, and the normal vector $\mathbf{n}$ will coincide with the unit vector $\mathbf{e}_r$.

According to (17A), the toroidal component of the field $\mathbf{E}_1(\hat{\mathbf{N}}_2\Psi_2)$ contains the only term $\sim\mathbf{e}_\phi$, and, consequently, relations (20A) and (19A) at $\boldsymbol{\tau} = \mathbf{e}_\Theta$ are identities at any values of the constants $D_n^{(2)}$ in solution (18A). From condition (19A) at $\boldsymbol{\tau} = \mathbf{e}_\phi$ and (17A) we obtain

$$r = R: \quad \frac{\partial\Psi_2}{\partial\Theta} = 0.$$

As this relation should be valid for any angle Θ, it is required that all constants $D_n^{(2)}$ in the solution for Ψ_2 be zero. Thus, the field $\mathbf{E}_1$ is completely defined by the scalar function Ψ_3

$$\mathbf{E}_1 = \hat{\mathbf{N}}_3\Psi_3 \equiv \nabla\times(\nabla\times\mathbf{r})\Psi_3$$

or, using (17A) and (18A), we obtain

$$\mathbf{E}_1 = \sum_n \exp(i\omega t)$$

$$\times\left\{E_{nr}n(n+1)\Im_n(\mu)\mathbf{e}_r + E_{n\Theta}\frac{\partial\Im_m(\mu)}{\partial\Theta}\mathbf{e}_\Theta\right\}, \quad (21A)$$

$$E_{nr} = D_n k^{-1/2} r^{-3/2} H^{(2)}_{n+1/2}(kr),$$

$$E_{n\Theta} = 2E_{nr} + r\frac{\partial E_{nr}}{\partial r}.$$

5. The total electrical field produced by the drop can be written in the form

$$\mathbf{E}_1 = -\nabla\frac{Q}{r} + \sum_n \exp(-i\omega t) \quad (22A)$$

$$\times\left\{E_{nr}n(n+1)\Im_n(\mu)\mathbf{e}_r + E_{n\Theta}\frac{\partial\Im_n(\mu)}{\partial\Theta}\mathbf{e}_\Theta\right\}.$$

To find the field strength near the drop surface, it is necessary to expand Hankel's function in a series for small values of its argument ($kr \ll 1$). According to [12], the second Hankel's function can be presented in the form

$$H^{(2)}_{n+1/2}(kr) = i^{n+1}\sqrt{\frac{2}{\pi z}}\exp(-iz)y_n\left(-\frac{1}{iz}\right),$$

where polynomial $y_n(\xi)$ is defined by Rodrigues' formula

$$y_n(\xi) = 2^{-n}\exp\left(\frac{2}{\xi}\right)\frac{d^n}{d\xi^n}\left[\xi^{2n}\exp\left(-\frac{2}{\xi}\right)\right].$$

Using these relations, it is not difficult to obtain the necessary expansion for Hankel's function about a small argument value

$$H^{(2)}_{n+1/2}(kr) = i\sqrt{\frac{2}{\pi}}\frac{(-1)^n C_n}{r^{n+1/2}}\left(\frac{c}{\omega}\right)^{n+1/2}\left(1 - i\frac{\omega}{c}r\right), \quad (23A)$$

where $C_n \equiv 1\times 3\times 5\times 7\times\ldots\times(2n+1)$.

By virtue of (23A) the expressions for E_{nr} and $E_{n\Theta}$ can be written in the following way

$$E_{nr} = D_n\frac{i(-1)^n C_n}{r^{n+2}}\left(\frac{c}{\omega}\right)^{n+1}\sqrt{\frac{2}{\pi}}\left(1 - i\frac{\omega}{c}r\right), \quad (24A)$$

$$E_{n\Theta} = -D_n\frac{i(-1)^n C_n n}{r^{n+2}}\left(\frac{c}{\omega}\right)^{n+1} \times\sqrt{\frac{2}{\pi}}\left[1 - ir\frac{\omega}{c}\frac{(n-1)}{n}\right]. \quad (25A)$$

To determine the unknown constant D_n we take into account that the surface of a conductive drop is equipotential; that is, the projection of the vector $\mathbf{E}$ onto the unit vector $\boldsymbol{\tau}$ tangential to the drop surface is zero

$$\boldsymbol{\tau}(\mathbf{r}, t)\mathbf{E}(\mathbf{r}, t) = 0. \quad (26A)$$

To find the unit vector tangential to the perturbed drop surface we first find a unit vector normal to it in the form

$$\mathbf{n} = \frac{\nabla F}{|\nabla F|},$$

where $F = r - R - \xi(\Theta, t) = 0$ is the equation of the perturbed drop surface.

Now we obtain

$$\mathbf{n} = \mathbf{e}_r - \frac{1}{r}\frac{\partial\xi}{\partial\Theta}\mathbf{e}_\Theta. \quad (27A)$$

Due to axial symmetry of the problem, the tangential unit vector in the latitudinal direction coincides with the corresponding unit vector of the spherical coordinate system $\boldsymbol{\tau}_\phi = \mathbf{e}_\phi$. The tangential unit vector in the meridional direction can be found using a known vector relation

$$\boldsymbol{\tau}_\Theta = \mathbf{n}\times\boldsymbol{\tau}_\phi. \quad (28A)$$

Using (27A), the expression for the tangential unit vector can be written in the form

$$\boldsymbol{\tau}_\Theta = -\frac{1}{r}\frac{\partial\xi}{\partial\Theta}\mathbf{e}_r - \mathbf{e}_\Theta. \quad (29A)$$

Having substituted (29A) and (22A) into (26A) we obtain a condition of equipotentiality of a conductive droplet

$$E_0\frac{1}{R}\frac{\partial\xi}{\partial\Theta} + E_{n\Theta}\frac{\partial\Im_n(\mu)}{\partial\Theta} = 0. \quad (30A)$$

Substitution of the expressions (8A) for E_0 and (25A) for $E_{n\Theta}$ into (30A) gives an expression for the unknown constant D_n

$$D_n = \frac{Q\alpha_n R^{n-1}}{in(-1)^n C_n\sqrt{\frac{2}{\pi}}}\left(\frac{\omega}{c}\right)^{n+1}\left[1 + iR\frac{\omega}{c}\frac{(n-1)}{n}\right]. \quad (31A)$$

Using (31A), (24A), and (25A) we find from (22A) an expression for the electrical field strength $\mathbf{E}(r, \Theta, t)$

$$\mathbf{E} = \frac{Q}{\varepsilon r^2}\mathbf{e}_r + \sum_n \exp(i\omega t)\left(\frac{R}{r}\right)^{n+1}\frac{Q\alpha_n}{R^3}$$

$$\times\left\{\left[1 + i\frac{\omega}{c}(R-r) - i\frac{\omega}{c}R\frac{1}{n}\right](n+1)\Im_n(\mu)\mathbf{e}_r\right.$$

$$\left. -\left[1 + i\frac{\omega}{c}(R-r)\frac{(n-1)}{n}\right]\frac{\partial\Im_n(\mu)}{\partial\Theta}\mathbf{e}_\Theta\right\}.$$

6. On the basis of the expression obtained one can find the pressure P_* exerted by the electrical field on the

droplet surface

$$r = R + \xi: \quad P_* = \frac{\varepsilon}{8\pi}(\mathbf{E})^2. \tag{32A}$$

The pressure determined in this way pertains to the perturbed droplet surface defined by $r = R + \xi$ whereas in hydrodynamics of the waves of infinitely small amplitude the dynamic boundary condition (9) is applied to the undisturbed droplet surface ($r = R$). Therefore we expand (32A) in the vicinity of $r = R$ in a series including terms up to the first order of smallness with respect to $|\xi|/R$

$$r = R: \quad P_* \approx \frac{\varepsilon}{8\pi}(\mathbf{e})^2 + (\boldsymbol{\xi}\nabla)\frac{\varepsilon}{8\pi}(\mathbf{E})^2.$$

Let us differentiate this expression once with respect to time (at $\Theta = \text{const}$) to find the time derivative of the pressure on the droplet surface produced by the electrical field of the droplet's own charge

$$r = R: \quad \frac{\partial}{\partial t}\left(\frac{\varepsilon}{8\pi}(\mathbf{E})^2\right) = \frac{i\omega}{4\pi}\frac{Q^2}{\varepsilon R^5}\sum_n \alpha_n \Im_n(\mu)$$
$$\times \exp(i\omega t)\left((n-1) - \frac{i\omega R}{c}\frac{(n+1)}{n}\right).$$

ACKNOWLEDGMENTS

The work was supported by a grant from the President of the Russian Federation (grant no. 00-15-9925).

REFERENCES

1. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
2. A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (5), 22 (2000) [Tech. Phys. **45**, 543 (2000)].
3. D. F. Belonozhko and A. I. Grigor'ev, Élektrokhim. Obrab. Met., No. 4, 17 (2000).
4. L. G. Kachurin, *Physical Principles of Influence on Atmospheric Processes* (Gidrometeoizdat, Leningrad, 1990).
5. V. I. Kalechits, I. E. Nakhutin, and P. P. Poluéktov, Dokl. Akad. Nauk SSSR **262**, 1344 (1982).
6. K. V. Beard, Rev. Geophys. **25**, 357 (1987).
7. V. V. Sterlyadkin, Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana **24**, 613 (1988).
8. K. V. Beard and Ali Tokay, Geophys. Res. Lett. **18**, 2257 (1991).
9. J. W. Schweizer and D. N. Hanson, J. Colloid Interface Sci. **35**, 417 (1971).
10. *Handbook of Clouds and Cloudy Atmosphere*, Ed. by I. P. Mazin and A. Kh. Khrgian (Gidrometeoizdat, Leningrad, 1989).
11. I. P. Mazin and S. M. Shmeter, *Clouds. Structure and Physics of Formation* (Gidrometeoizdat, Leningrad, 1983).
12. A. F. Nikiforov and V. B. Uvarov, *Special Functions of Mathematical Physics* (Nauka, Moscow, 1984; Birkhauser, Basel, 1987).

Translated by N. Mende

Technical Physics, Vol. 47, No. 1, 2002, pp. 64–69. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 66–71.

SOLID-STATE ELECTRONICS

Electron Transport in GaAs Nanostructures under Irradiation

N. V. Demarina and S. V. Obolensky

Lobachevsky State University, pr. Gagarina 23, Nizhni Novgorod, 603600 Russia

e-mail: obolensk@rf.unn.runnet.ru

Received February 5, 2001

Abstract—The results of theoretical and experimental investigations into the *n*-GaAs electrophysical characteristics under irradiation are reported. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

To date, a great amount of experimental data for the mobility, concentration, and lifetime of charge carriers in the basic semiconductors (Si, Ge, and GaAs) irradiated by ions, neutrons, and electrons have been amassed. As the sizes of the solid structures are now approaching the nanometer range, the effects caused by the nonequilibrium state and the heterogeneity of the electron–hole gas (for example, drift velocity burst) become more and more important. The burst of drift velocity is due to the fact that an electron, moving in a nanostructure for a time comparable to the energy or momentum relaxation time, reaches the region of the strong field and gains the energy more than one order of magnitude higher than the thermal energy. Of interest is electron transport in the structures with nanoclusters of radiation-induced defects (disordered regions) when the sizes of the clusters are comparable to both the sizes of the structure and the electron wavelength. A correlation between the relaxation times of the electron energy and momentum and the irradiation fluence under electric fields up to 100 kV/cm should be found.

MATHEMATICAL MODEL

The Boltzmann kinetic equation is usually used to describe the electron–hole gas parameters in a semiconductor. The numerical solution to the Boltzmann equation using the Monte Carlo method allows one to carefully study the interaction of the electron gas with the semiconductor crystal lattice, to take into account the dispersion pattern (the presence of many valleys and their nonparabolicity), and to consider almost all mechanisms of charge carrier scattering in a wide range of electric fields [1].

It is well known [2, 3] that the presence of charged scattering centers, such as dopant ions, substantially influences the electrophysical parameters of semiconductors. Therefore, one should consider the effect of charged radiation-induced defects when solving the problem stated especially if irradiation doses are such that the defect concentration is higher than that of dopant ions. Of no less importance are large clusters of point defects—disordered regions (DRs)—produced by fast heavy particles. The DRs act as barriers for the charged particle flow, radically changing the kinetic parameters.

In irradiation experiments, neutron radiation is the most convenient to use. Neutrons bear no charge and, hence, offer high penetrability. Therefore, the defect distribution in a homogeneous material is uniform and easy to calculate. Since neutron radiation produces not only point defects (PD) but also DRs, a low-energy proton radiation was used to evaluate the PD effect.

In the model, radiation-induced defects were taken into account by introducing additional mechanisms of carrier scattering by PDs and DRs. The frequency of scattering by charged PDs was calculated in the Born approximation for the screened Coulomb potential [4]:

$$\lambda_i(k) = \frac{2\sqrt{2}\pi N_{pd} q^4 m^{*1/2}}{\varepsilon^2 \hbar^2 \beta^2 (N_{pd})} \frac{(1+2\alpha W)}{[W(1+\alpha W)]^{1/2}},$$

$$\alpha = \frac{1}{E_g}\left(1 - \frac{m^*}{m_0}\right)^2, \quad \beta^2 = \frac{4\pi n q^2}{\varepsilon k_B T}.$$

Here, $n(N_{pd})$ is the electron concentration dependent on the concentration of traps; N_{pd} is the charged point defect concentration; q is the electron charge; T is the lattice temperature; k_B is the Boltzmann constant; h is the Planck constant; ε is the permittivity; E_g is the band gap; and m^* and m_0 are the effective mass in a valley and the free-electron mass, respectively.

The point defect concentration in the material irradiated is determined as [5]

$$N_{pd} = G_e F_e,$$

where $G_e = 5$ cm^{-1} is the rate of PD generation, N_{pd} is the PD concentration (cm^{-3}), and F_e is the electron fluence (cm^{-2}).

A change in the screening length in the material irradiated because of the capture of free carriers by radiation-induced traps was also taken into account.

The disordered regions were described within the Gossik model [2, 6]. According to is model, the interior of a DR contains a neutral core enriched with divacancies and surrounded by a charged shell consisting of vacancy–impurity and vacancy–defect complexes. The entire DR is inside of the space charge region whose size depends on the carrier concentration in the DR-free part of the crystal. The shape of the disordered region can approximately be considered as spherical. In the frame of the Gossik model, the accurate determination of the DR parameters became possible with invoking experimental data on the ultimate position of the Fermi level in irradiated GaAs, as well as on the concentration and energy levels of radiation-induced defects [7].

The radii of the damaged region, r_1; inner central core, r_0; and space charge layer, r_2, are obtained from the solution of the Poisson equation, which relates the potential φ in the space charge layer of the imperfect region and the potential in the surrounding "perfect" part of the semiconductor. In spherical coordinates, it is written as

$$\frac{d}{dr}\left(r^2\frac{d\varphi}{dr}\right) = -\frac{4\pi}{\varepsilon}r^2Q(r).$$

Here, r is the radius vector, $Q(r)$ is the charge density, and ε is the permittivity of the material.

In the general case, the right of the equation includes the charge of both fixed centers and mobile carriers. Neutron irradiation pins the Fermi level near the level E_v + 0.6 eV at the center of the disordered region [8], which is the boundary condition for the Poisson equation.

In the presence of the defects of both types, the DR concentration N_{dr} and the PD concentration N_{pd} as a function of the neutron fluence are given by [2]

$$N_{pd} = G_{pd\,n}F_n, \quad N_{dr} = K_t(F_n)F_n,$$

$$K_t(F_n) = 44.5N_d^{-0.55}, \quad N_1 = 7.9N_d^{-0.39}.$$

Here, $G_{pd\,n}$ = 50 cm^{-1}; $G_{dr\,n}$ = 0.2 cm^{-1}; N_{pd} and N_{dr} are the concentrations of point defects and disordered regions in cm^{-3}, respectively; N_d is the ionized dopant concentration in 10^{15} cm^{-3}; N_1 is the concentration of charged centers in a DR; and F_n is the neutron fluence in cm^{-2}.

The classical movement of the particles over the large-scale potential relief, appearing near the disordered region [6], allows us to consider carrier scattering by DRs as scattering by inclusions nontransparent for the particles. The scattering frequency was determined through the inclusion crosssection dependent on the charge carrier energy:

$$\Sigma_{dr} = 2\pi r_{ef}^2(kr_{ef} \gg 1) \quad \text{and} \quad \Sigma_{dr} = 4\pi r_{ef}^2(kr_{ef} \ll 1),$$

where k is the electron wave vector.

In this case, the scattering frequency for the disordered region in the jth valley takes the form

$$\lambda_j(W) = v_j\Sigma_{dr}N_{dr},$$

where v_j is the electron velocity in the jth valley.

It should be noted that this type of scattering, like that by charged PDs, is elastic; however, scattering by PDs is small-angle, while that by DRs is randomizing.

EXPERIMENT

We studied MESFETs with a built-in channel, which are convenient for taking the electrophysical characteristics of semiconductor layers. The *I–V* and *C–V* characteristics of the structure provide information on the mobility and dopant profiles in the channel and in the buffer conducting layer under the gate. Varying the drain voltage, we obtained the field dependence of the charge carrier velocity, while varying the gate voltage, we made measurements in a wide range of dopant concentrations (10^{15}–10^{18} cm^{-3}) at the channel–buffer interface. The relaxation times of the charge carrier momentum, as well as the nature of scattering processes, were judged from the *I–V* characteristics of the nanometer-gate transistors. The well-developed theory of charge carrier transfer in MESPETs [9] allows us to convert the measurements to the material parameters with a required accuracy.

In this work, conventional n^+–n–n^--GaAs structures with different dopant concentrations in the n-layer were used as initial semiconductor structures. Transistors with a gate length of 0.03 to 1.5 μm were fabricated on each of the structures. The gate width was varied from 50 to 500 μm. The transistors with a gate length of more than 0.25 μm were fabricated by the standard technology [9]. The devices with an effective gate length of 30 nm were made by anisotropic etching. This method allows one to etch a *V* groove in the GaAs active layer and to metallize the gate by the self-aligned method [10].

To study the PD effect on carrier transport in the submicron structures, we took the *I–V* and *C–V* characteristics of the devices before and after proton irradiation with an energy of 30–90 keV and a dose of 0.01–0.3 μC. A Vezuviĭ implanter was used as a radiation source in the experiments. The effect of the DRs was studied with the *I–V* and *C–V* characteristics recorded before and after the irradiation by fission neutrons with an average energy of 1 MeV in a VIR-2M reactor. The irradiation adversely affects the transistor performance, as demonstrated by the dependences of the drain current, power gain coefficient, and noise factor on the neutron fluence.

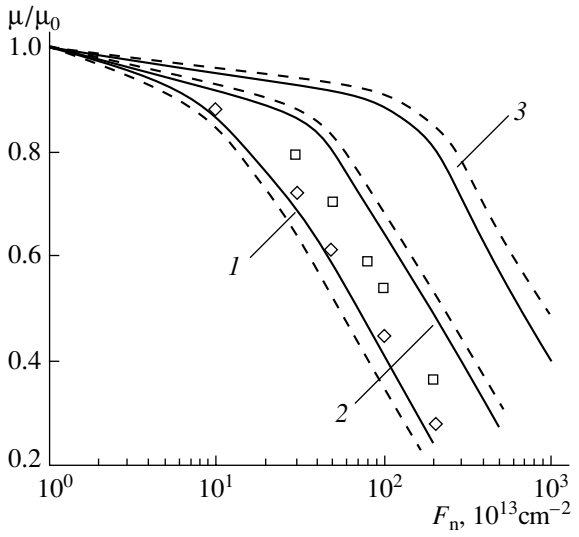


Fig. 1. Relative mobility μ/μ_0 vs. neutron fluence F_n. The initial dopant concentration is (*1*) 10^{15}, (*2*) 10^{16}, and (*3*) 10^{17} cm^{-3}. Solid curves, Monte Carlo calculation; dashed curves, experimental data; and symbols, analytical calculations ($\diamond$ 10^{15} and $\square$ 10^{16} cm^{-3}).

The concentration and the mobility of charge carriers in the transistor channel, as well as in the buffer layer adjacent to the channel, were found from the *I–V* and *C–V* characteristics at small drain voltages using the standard technique. Applying different voltages to the gate, we traced the variation of the *n*-GaAs characteristics in the range of dopant concentrations from 5×10^{15} (buffer layer) to 5×10^{17} cm^{-3} (channel) with an accuracy of 10–30% [11, 12]. It was found that the concentration and the mobility are well described by the relations [13, 14]

$$n = n_0(1 - \alpha F_n), \quad \alpha = 7.2 \times 10^{-4} n_0^{-0.77},$$
$$\mu_n = \mu_{n0}(1 - bF_n)^{-1}, \quad b = 7.8 \times 10^{-6} n_0^{-0.64}. \qquad (1)$$

Here, n_0 (cm^{-3}) is the initial concentration of free electrons, n (cm^{-3}) is the postirradiation electron concentration, F_n (cm^{-2}) is the neutron fluence, μ_{n0} (cm^2/(V s)) is the initial electron mobility, and μ_n is the postirradiation electron mobility.

The experimental fluence dependences of the mobility are shown in Fig. 1 (dashed curves).

The field dependences of the electron velocity at high electric fields were determined by measuring the transistor residual resistance [9]. The data points in Fig. 2 agree with the theoretical curves.

The momentum relaxation time is more sensitive to the presence of radiation-induced defects in the material in comparison with the energy relaxation time. This fact can be explained by different mechanisms responsible for relaxation. The electromagnetic character of the interaction between charge carriers and charged scattering centers provides energy conservation at collisions; hence, the degradation of the energy relaxation time cannot be detected experimentally.

To experimentally find the field dependence of the momentum relaxation time after irradiation, the *I–V* characteristics of the transistors with the 30-nm gate were recorded. In this case, the drain current is defined only by electron drift in the high field beneath the gate, while the diffusion current component is insignificant. Then, the momentum relaxation time τ_m is determined from the relationship

$$\tau_m(W(E)) = \frac{M(E)v_{dr}(E)}{qE} = \frac{m(E)I_d}{q^2EnS},$$

where m is the effective electron mass, E is the electric field intensity, q is the electron charge, n is the electron concentration, v_{dr} is the drift velocity, I_d is the drain current, and S is the channel cross-sectional area.

The experimentally found dependences of the momentum relaxation time on the electric field intensity agree with the theoretical ones (Figs. 3, 4).

ANALYTICAL RESULTS VERSUS EXPERIMENTAL DATA

Point Defects

The numerical calculations were preceded by the analytical evaluations of the radiation-induced defect concentration and the defect-concentration-dependent carrier mobility. The PD concentration was determined using the Khinchin–Piez cascade model [3] and the experimental data on the energy level position and the charge states of the defects [5]. The electron mobility in the *n*-GaAs was evaluated with the Bruks–Herring relationship [9]. For the high-resistivity material (dopant concentration $N_d = 10^{15}$ cm^{-3}), the twofold decrease in the mobility occurs when the PD concentration is one order of magnitude higher than N_d. For the low-resistivity GaAs ($N_d = 10^{17}$ cm^{-3}), this takes place when the PD concentration is comparable to N_d. Such behavior is associated with the weak dependence of the mobility on the concentration of charged scattering centers if this concentration is less than 10^{16} cm^{-3}.

The effects of charged PDs and dopant ions on the GaAs characteristics were compared using the Monte Carlo method. It was found that an increase in the dopant concentration in the material changes the characteristics because of a change in the angular distribution of scattering. In the case of scattering by charged point defects, as their concentration increases, scattering remains small-angle, while the scattering frequency increases, since the carrier concentration changes because of the capture of the carriers by radiation-induced traps. That is why the GaAs electrophysical characteristics, such as mobility, drift velocity, and relaxation times of the electron energy and momentum (Figs. 1–4) in the material with a dopant concentration

of 5×10^{16} cm^{-3} and in the undoped material with the same defect concentration differ by more than 20% at an electric field intensity $E < 10$ kV/cm.

At fields $E < 10$ kV/cm, scattering by charged PDs results in a considerable decrease in the mobility, drift velocity, and momentum relaxation time (twofold increase at defect concentrations comparable to the dopant concentration) and in an increase in the energy relaxation time (by 30%). At moderate field intensities, the effect of PD scattering is reduced (relative changes in the parameters are 20% for $E = 20$ kV/cm) and becomes inappreciable for $E = 100$ kV/cm.

Point Defects and Disordered Regions

We analytically compared the effect of charged PDs with that of DRs on radiation-induced changes in the *n*-GaAs mobility. It can be noted that PDs, whose concentration is two orders of magnitude higher than that of DRs, affect the mobility in the same way as DRs but only when the changes are small (less than 25%). When the relative changes in the mobility are large, the effect of the DRs dominates, causing the discrepancy between the experimental data and the calculations by Eq. (1). Equation (1) coincides well with the experimental results for relative changes of 25–30% and can be used as a rough estimate up to relative changes of about 50%, which is contradictory to the conclusions made in [13]. This may be explained by the fact that our epitaxial GaAs structures had a somewhat larger concentration of growth defects than the structures considered in [13].

The effect of the PDs and that of the DRs on the electrophysical characteristics of the material with different doping levels were compared with the Monte Carlo method. Since the size of the cluster space charge region depends on the carrier concentration, the radiation fluences were chosen so that the total volume of the space charge regions in both samples was equal. We took into account here that the carrier concentration decreases during irradiation and that the concentration of charged PDs in heavily doped samples is about one order of magnitude smaller than in low-doped ones [7]. Such conditions allow us to compare the relative effects of the PDs and DRs when their absolute concentrations in heavily doped and low-doped samples differ significantly.

At a small dopant concentration (10^{15} cm^{-3}), the disordered regions make a contribution to a change in the parameters comparable with that of the point defects only for high irradiation doses, while at a dopant concentration of 10^{17} cm^{-3}, the effect of the DRs dominates. This is because the concentration of radiation-induced PDs in the low-doped semiconductor is much higher than the dopant concentration. For the heavily doped samples, it is comparable to or even less than the dopant concentration. Although in the low-doped and in the heavily doped semiconductor the total volume of the high-resistivity regions is the same, the number of these regions in the latter case is larger, so that the electron–DR collision frequency grows. That is why the effect on the DRs on the conductivity dominates in the heavily doped samples.

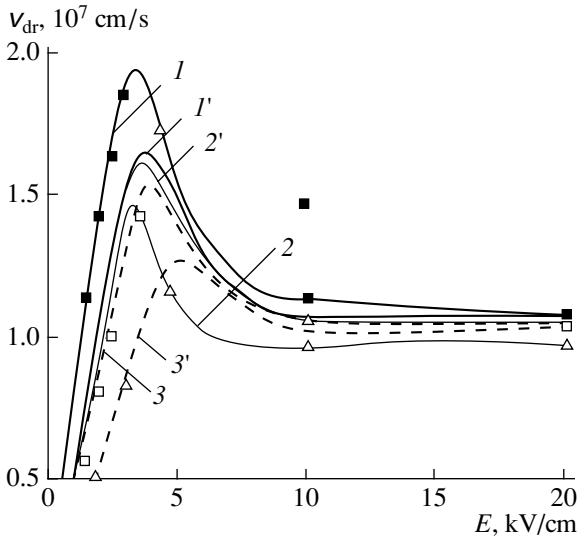


Fig. 2. Electron drift velocity v_{dr} in *n*-GaAs vs. electric field intensity E (*1*) without and with defects induced by (*2*) proton and (*3*) neutron irradiations. (*2*) $N_{pd} = 2.5 \times 10^{16}$ cm^{-3}; (*3*) $N_{pd} = 2.5 \times 10^{16}$ cm^{-3}, $N_{dr} = 10^{14}$ cm^{-3}; (*2*') $N_{pd} = 2 \times 10^{16}$ cm^{-3}; and (*3*') $N_{pd} = 2 \times 10^{16}$ cm^{-3}, $N_{dr} = 10^{15}$ cm^{-3}. The initial dopant concentration is (*1*–*3*) 10^{15} and (*1*'–*3*') 10^{17} cm^{-3}. Curves, Monte Carlo calculation; symbols, data points.

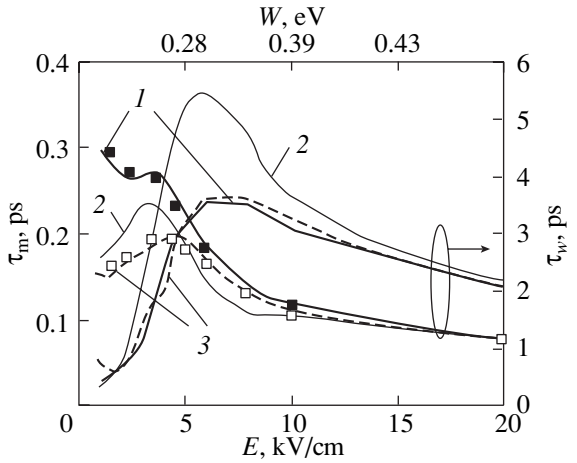


Fig. 3. Relaxation times of the electron momentum, τ_m, and energy, τ_w, in *n*-GaAs vs. electric field intensity E (*1*) without and with defects produced by (*2*) proton and (*3*) neutron irradiations. (*2*) $N_{pd} = 2.5 \times 10^{16}$ cm^{-3} and (*3*) $N_{pd} = 2.5 \times 10^{16}$ cm^{-3}, $N_{dr} = 10^{14}$cm^{-3}. The initial dopant concentration is 10^{15} cm^{-3}. Curves, Monte Carlo calculation; symbols, data points.

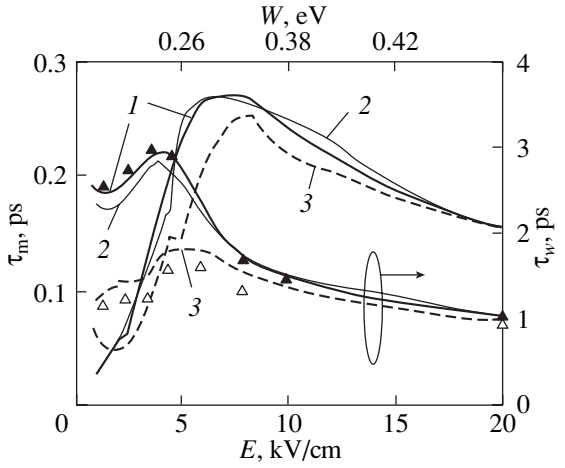


Fig. 4. The same as in Fig. 3 for (*2*) $N_{pd} = 2 \times 10^{16}$ cm^{-3} and (*3*) $N_{pd} = 2 \times 10^{16}$ cm^{-3}, $N_{dr} = 10^{15}$ cm^{-3}. The initial dopant concentration is 10^{17} cm^{-3}.

The field dependence of the electron velocity under neutron irradiation is well approximated by the formulas

$$v_{dr}(E) = \frac{\mu_n(F_n) + 2v_s\left(a(F_n)\dfrac{E}{1.03E_0(F_n)}\right)^4}{1 + 2a(F_n)(E/E_0(F_n))^{3.99}},$$

$$v_s = v_{s0}E^m, \quad v_{s0} = 0.45\eta + 1.6559,$$

$$m = -0.08\eta - 0.17,$$

$$E_0(F_n) = E_0^{(1)}\xi^2 + E_0^{(2)}\xi + E_0^{(3)},$$

$$a(F_n) = a^{(1)}\xi^2 + a^{(2)}\xi + a^{(3)},$$

$$E_0^{(1)} = 0.001\exp(3.18\eta),$$

$$E_0^{(2)} = 0.019\exp(3.61\eta),$$

$$E_0^{(3)} = 3.951\exp(-0.01\eta),$$

$$a^{(1)} = 0.0003\exp(6.14\eta),$$

$$a^{(2)} = 0.063\exp(-0.41\eta),$$

$$a^{(3)} = 0.72\exp(-0.33\eta).$$

Here, E is the electric field intensity in kV/cm, $E_0 = 4$ kV/cm, $\eta = N_d/N_{d0}$, N_d is the dopant concentration, $N_{d0} = 10^{17}$ cm^{-3}, $\xi = F_n/F_{n0}$, F_n is the neutron fluence, and $F_{n0} = 10^{14}$cm^{-2}.

Analytical expressions approximating the energy dependence of the momentum relaxation time are as follows:

$$\tau_m(W, F_n) = a_0W^4 + a_1W^3 + a_2W^2 + a_3W + a_4, \quad a_i = a_i^{(1)}\xi + a_i^{(2)},$$

$$a_0^{(1)} = -0.03\eta + 0.1266, \quad a_0^{(2)} = -0.89\eta - 3.8688,$$

$$a_1^{(1)} = -0.08\eta - 0.2478, \quad a_1^{(2)} = 1.3\eta + 8.9344,$$

$$a_2^{(1)} = -0.03\eta + 0.1307, \quad a_2^{(2)} = -0.78\eta - 6.4773,$$

$$a_3^{(1)} = -0.03\eta + 0.0097, \quad a_3^{(2)} = 0.5\eta + 1.0427,$$

$$a_4^{(1)} = 0.01\eta - 0.0111, \quad a_4^{(2)} = -0.08\eta + 0.2006.$$

Here, $\xi = F_n/F_{n0}$, $\eta = N_d/N_{d0}$, F_n is the neutron fluence, $F_{n0} = 10^{14}$ cm^{-2}, W is the electron energy in eV, N_d is the dopant concentration, and $N_{d0} = 10^{17}$ cm^{-3}.

CONCLUSION

Our calculations and experiments showed that in *n*-GaAs radiation-induced defects greatly influence the transport of charge carrier with energies less than 0.5 eV, while at high energies, optical phonon scattering and intervalley scattering play a decisive role. The results obtained may be used for designing radiation-resistant high-frequency semiconductor devices, specifically hot-charge-carrier devices.

ACKNOWLEDGMENTS

The authors thank M.A. Kitaev and A.G. Fefelov for the submission of the experimental data and also V.D. Skupov, V.K. Kiselev, A.N. Kachemtsev, V.A. Kozlov, M.N. Drozdov, and D.I. Tetel'baum for the valuable discussions.

The work was partially supported by the grant SfP-973799 and by the Program "Physics of Solid-State Nanostructures" (grant no. 99-1142).

REFERENCES

1. W. Fawsett, D. A. Boardman, and S. J. Swain, Phys. Chem. Solids **31**, 1963 (1970).
2. V. L. Vinetskiĭ and G. A. Kholodar', *Radiation Physics of Semiconductors* (Naukova Dumka, Kiev, 1979).
3. V. S. Vavilov, N. P. Kekelidze, and L. S. Smirnov, *Interaction of Radiation with Semiconductors* (Nauka, Moscow, 1988).
4. J. G. Ruch and W. J. Fawsett, Appl. Phys. **41**, 3843 (1970).
5. D. Lang, in *Radiation Effects in Semiconductors: Invited and Contributed Papers from the International Conference on Radiation Effects in Semiconductors, Dubrovnik*, Ed. by N. B. Urli and J. W. Corbett (Inst. of Physics, Bristol, 1977).

6. R. F. Konopleva, V. L. Pitvinov, and N. A. Ukhin, *Radiation-Induced Damage to Semiconductors by High-Energy Particles* (Moscow, 1971).
7. V. P. Klad'ko and V. P. Plyatsko, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 261 (1998) [Semiconductors **32**, 235 (1998)].
8. V. N. Brudnyi, S. N. Grinyaev, and V. E. Stepanov, Physica B: Condens. Matter (Amsterdam) **212**, 429 (1995).
9. M. Shur, *GaAs Devices and Circuits* (Plenum, New York, 1987; Mir, Moscow, 1991).
10. S. V. Obolensky and M. A. Kitaev, Pis'ma Zh. Tekh. Fiz. **26** (10), 13 (2000) [Tech. Phys. Lett. **26**, 408 (2000)].
11. S. V. Obolensky and G. P. Pavlov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 413 (1995) [Semiconductors **29**, 211 (1995)].
12. S. V. Obolensky and M. A. Kitaev, Mikroélektronika **30**, 7 (2001).
13. *VLSI Electronics: Microstructure Science*, Vol. 11: *GaAs in Microelectronics*, Ed. by N. G. Einspruch and U. Uissmen (Academic, Orlando, 1985; Mir, Moscow, 1988).
14. N. V. Demarina and S. V. Obolensky, Microelectron. Reliab. **39**, 1247 (1999).

Translated by M. Astrov

Technical Physics, Vol. 47, No. 1, 2002, pp. 70–73. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 72–75.

OPTICS, QUANTUM ELECTRONICS

The Effect of Thermionic Emission on the Absorption of Ultrashort Laser Pulses in Semiconductors

P. V. Lobzenko, N. A. Evtushenko, V. A. Novikov, and R. G. Irishin

Rostov Military Institute of Missile Forces, Rostov-on-Don, 344027 Russia

Received April 12, 2001

Abstract—The interaction between high-power ultrashort laser pulses and the semiconductor surface is considered with allowance for the effect of thermionic emission on the temperature dynamics of the electron and ion subsystems in a semiconductor medium. The parameters of the pulses for which thermionic emission must be taken into account are determined. A computational scheme that makes it possible to solve the problem in the three-dimensional formulation is obtained. © *2002 MAIK "Nauka/Interperiodica".*

The investigation of physical processes related to the absorption of light by semiconductor structures is stimulated by the extensive development of semiconductor sensors, which are widely used in aviation, rocket production, robotics, etc. To date, the problem of absorption of high-power short (10^{-9} s) and ultrashort (below 10^{-12} s) optical pulses by semiconductors [1–4] is a most interesting and, at the same time, little studied line of inquiry in this field.

The mathematical models describing laser-irradiated semiconductor structures usually use a number of assumptions that simplify the solution to the problem. For example, it is often assumed that the overheating of the electron subsystem relative to the phonon one [2, 3] is negligible. At the same time, for laser pulses shorter than the thermalization time τ_t (the characteristic time of electron–lattice energy exchange), the process of semiconductor heating changes both qualitatively and quantitatively. This feature of short-pulse absorption in semiconductors has been noted, for example, in [1, 4]. For example, in the case of ultrashort pulses [$\tau_p \le 0.1$–1.0×10^{-12} s], when $\tau_p \ll \tau_t$, the radiation absorption by the electron subsystem and the energy exchange with the phonon subsystem are significantly affected by thermionic emission. This results in a change in the number of free charge carriers and their contribution to the electron, T_e, and ion, T_i, temperatures and, hence, to the total temperature of the system. In this study, we consider the absorption of ultrashort laser pulses by a semiconductor and determine the temperatures of its electron and ion subsystems with regard for thermionic emission.

To describe this process in semiconductors, we use a set of equations [2, 4],

$$\frac{\partial n}{\partial t} = D\Delta n - \frac{n}{\tau} + P_1,$$

$$\frac{\partial T_e}{\partial t} = \frac{\lambda_e}{C_e}\Delta T_e + \frac{(1-\beta)E_g}{C_e\tau}(n-n_t) - \frac{G}{C_e}(T_e - T_i) + \frac{1}{C_e}P_2, \quad (1)$$

$$\frac{\partial T_i}{\partial t} = \frac{\lambda_i}{C_i}\Delta T_i + \frac{\beta E_g}{C_i\tau}(n-n_t) + \frac{G}{C_i}(T_e - T_i).$$

Here, n is the concentration of laser-induced nonequilibrium carriers; n_t is the concentration of free carriers responsible for the thermionic current; C_e, C_i and λ_e, λ_i are the electronic and ionic specific heats and thermal conductivities, respectively; G is the electron–lattice energy exchange coefficient; E_g is the forbidden gap of the semiconductor; τ is the recombination time of the nonequilibrium carriers; β is a portion of the energy released in the recombination process that is acquired by the lattice; D is the diffusion coefficient of the carriers; $P_1 = \alpha(1-R)\lambda_p/(\vartheta_{ph}h)I\exp\{-\alpha x\}$; $P_2 = \alpha(1-R)(\vartheta_{ph}h/\lambda_p - E_g)\lambda_p/(\vartheta_{ph}h)I\exp\{-\alpha x\}$, $I \ne 0$ at $t \le \tau_p$; R and α are the radiation reflection and absorption coefficients of the material; I, λ_p, and τ_p are the pulse intensity, wavelength, and duration; ϑ_{ph} is the speed of light in the material; and h is the Planck constant.

We choose boundary conditions that make it possible to simplify the analysis. It is assumed that the concentration of the carriers, the subsystem temperatures, and the temperature gradients outside and inside the semiconductor vanish far away from the region where the radiation is absorbed:

$$-\lambda_e\frac{\partial T_e}{\partial x} = 0, \quad -D\frac{\partial n}{\partial x} = 0 \quad \text{at} \quad x = 0, \infty, y, z, t. \quad (2)$$

The initial conditions correspond to room temperature:

$$n(\mathbf{r}, t) = n_0, \quad T_e(\mathbf{r}, t) = T_i(\mathbf{r}, t) = T_0 \quad \text{at} \quad t \le 0, \quad (3)$$

where $n_0 = 10^{21}$ m^{-3} and $T_0 = 300$ K.

1063-7842/02/4701-0070$22.00

The quantity n_t entering into (1) and distinguishing our model from the available ones is determined as follows. Expressing the thermionic current [5] over the surface of the semiconductor with the electrons heated to a temperature T_e through their flux n_t moving with an average velocity v_m, we have

$$n_t = B/(e v_m) T_e^2 \exp\{-\varphi/kT_e\}, \quad (4)$$

where $B = 4\pi m_e e k^2/h^3$; φ is the work function; k is the Boltzmann constant; and ϑ_m, m_e, and e are the average velocity of the electrons, the elementary mass, and the elementary charge, respectively.

The average velocity, in its turn, depends on the electromagnetic field strength E, which is related to the radiation intensity I [6] and the carrier mobility μ [7]:

$$v_m = \mu E = De/(kT_e)19\sqrt{I}. \quad (5)$$

Substituting this relationship into (4) and taking into account the numerical value of B, we come to the expression for the number of the electrons responsible for the thermionic current and making no contribution to the semiconductor temperature:

$$n_t = \frac{4\pi m_e}{De}\left(\frac{k}{h}\right)^3 \frac{T_e^3}{\sqrt{I}} \exp\{-\varphi/(kT_e)\}. \quad (6)$$

To circumvent difficulties in taking the awkward integrals and to avoid errors in the Fourier transforms [3] (which eventually were obtained numerically!), we solved the set of equations (1) by the finite-difference method [8]. The computational scheme has the form

$$n_{jkm}^{i+1} = C_1 \Lambda_{xyz} n_{jkm}^{i} + C_2 n_{jkm}^{i} + \Delta t\, f_1^{i+0.5},$$

$$T_{E_{jkm}}^{i+1} = C_3 \Lambda_{xyz} T_{E_{jkm}}^{i}$$

$$+ C_4[n_{jkm}^{i} - C_5 (T_{E_{jkm}}^{i})^3 \exp\{-\varphi/(kT_{E_{jkm}}^{i})\}]$$

$$- C_6 (T_{E_{jkm}}^{i} - T_{L_{jkm}}^{i}) + C_7 f_2^{i+0.5} + T_{E_{jkm}}^{i}, \quad (7)$$

$$T_{L_{jkm}}^{i+1} = C_8 \Lambda_{xyz} T_{L_{jkm}}^{i}$$

$$+ C_9[n_{jkm}^{i} - C_5 (T_{E_{jkm}}^{i})^3 \exp\{-\varphi/(kT_{E_{jkm}}^{i})\}]$$

$$+ C_{10}(T_{E_{jkm}}^{i} - T_{L_{jkm}}^{i}) + T_{E_{jkm}}^{i}.$$

Here, $C_1 = D\Delta t/\Delta h^2$; $C_2 = 1 - \Delta t/\tau$; $C_3 = \lambda_e \Delta t/(C_e \Delta h^2)$; $C_4 = (1 - \beta) E_g \Delta t/(C_e \tau)$; $C_5 = 4\pi m_e k^3/(eD\sqrt{I}\, h^3)$; $C_6 = \Delta t G/C_e$; $C_7 = \Delta t/C_e$; $C_8 = \lambda_L \Delta t/(C_L \Delta h^2)$; $C_9 = \beta E_g \Delta t/(C_L \tau)$; $C_{10} = G\Delta t/C_L$; $\Lambda_{xyz} T = (\Lambda_x + \Lambda_y + \Lambda_z) T$ is the sum of the second-order difference derivatives $\Lambda_x T = (T_{j-1km}^{i} - 2T_{jkm}^{i} + T_{j+1km}^{i})/\Delta h_x^2$, $\Lambda_y T = (T_{jk-1m}^{i} - 2T_{jkm}^{i} + T_{jk+1m}^{i})/\Delta h_y^2$, and $\Lambda_z T = (T_{jkm-1}^{i} - 2T_{jkm}^{i} + T_{jkm+1}^{i})/\Delta h_z^2$; $\Delta h_x = \Delta h_y = \Delta h_z = \Delta h$ are the coordinate increments; Δt is the time increment; $f_1^{i+0.5}$ and $f_2^{i+0.5}$ are the difference approximations of the functions P_1 and P_2, respectively, entering in relationship (1) for the time instant $t_{i+0.5} = t_i + 0.5\Delta t$; and T_E and T_L are the difference electron and lattice (ion) temperatures, respectively.

Table

Action no.	$\tau_p \times 10^{-12}$, s	$Q_p \times 10^{-2}$, J	$I_0 \times 10^9$, W/cm^2
1	70.0	89.3	5.1
2	10.0	8.9	35.6
3	1.0	8.7	348.0
4	0.1	7.8	3120.0

The computational stability of scheme (7) is provided by satisfying the condition [8]

$$\Delta t/\Delta h^2 \le 0.5. \quad (8)$$

Note also that the set of equations (1) is nonlinear, and its coefficients are temperature-dependent. The most significant feature of our scheme is taking into account the temperature dependence of the electron specific heat, which is given by [9]

$$C_e = \pi^2/2nk^2 T_e/Q_F, \quad (9)$$

where Q_F is the Fermi energy.

From relationship (6), it follows that the number of the electrons responsible for the thermionic current is determined in the main by their temperature and work function. Apparently, for n_t to be as high as possible, it is necessary that the energy of the laser pulses Q_p be spent on heating the electron gas provided that the electron–lattice energy exchange is insignificant. There-

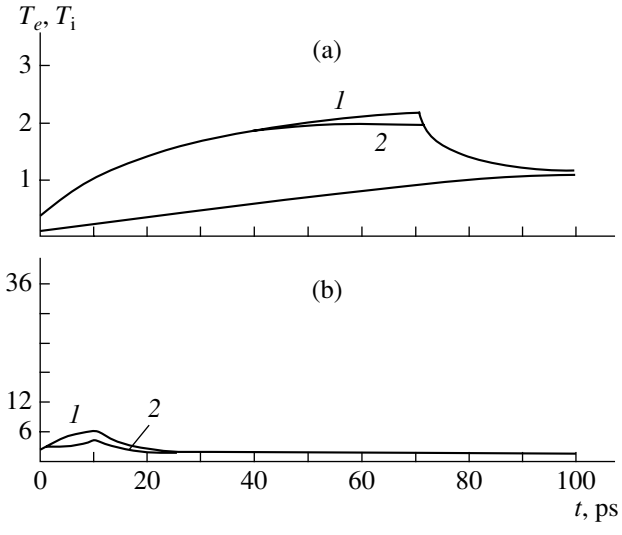


Fig. 1. Time dependence of the electron temperature under the action of ultrashort laser pulses: (a, b) laser actions 1 and 2 (see table), respectively (the irradiation depth is $x = 0.45 \times 10^{-6}$ m).

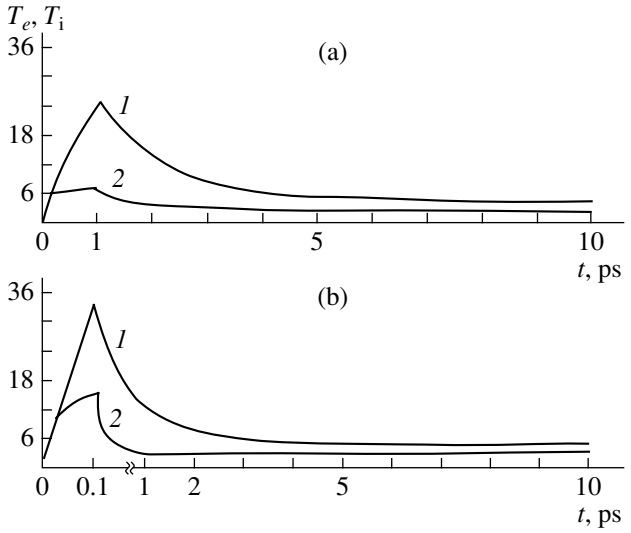


Fig. 2. The same as in Fig. 1 for laser actions (a) 3 and (b) 4.

fore, the irradiation conditions (see table) were chosen so that the sample began to melt at the end of the characteristic time of energy exchange between the electron and ion subsystems, i.e., at the end of the thermalization time $\tau_t = C_i/G$ (the area exposed was fixed: $S_{fix} = 25 \times 10^{-6}$ m^2).

The evolution of the heating process for the silicon sample is shown in Figs. 1–3. Curves *1*, *3* and *2*, *4* correspond to the electron and ion temperatures calculated without and with allowance for thermionic emission, respectively. For convenience, the values of T_e and T_i are normalized to the melting point of the sample, $T_{mel}^{Si} = 1700$ K.

The effect of thermionic emission on the heating process was estimated as the percentage difference in the sample temperatures with, T_i^{em}, and without, T_i, the emission at the instant of melting:

$$\delta = (T_i - T_i^{em})/T_i^{em} \times 100\%. \qquad (10)$$

As is seen from Fig. 3, the effect of thermionic emission must be taken into account under the above conditions even at $\tau_p \le 10^{-11}$ s = τ_r (the recombination time of the nonequilibrium carriers). This is explained by the fact that the electrons produced by the irradiation, when escaping the material, carry away a significant portion of the energy rather than transfer it to other electrons and to the lattice, thereby heating the sample. The significant decrease in the temperature of the electron subsystem when the exposure time decreases in the presence of the emission is accounted for by the same reason (Figs. 1b, 2).

The effect of the emission is the most pronounced when the pulse duration is less than the characteristic time of energy exchange in the electron subsystem, i.e.,

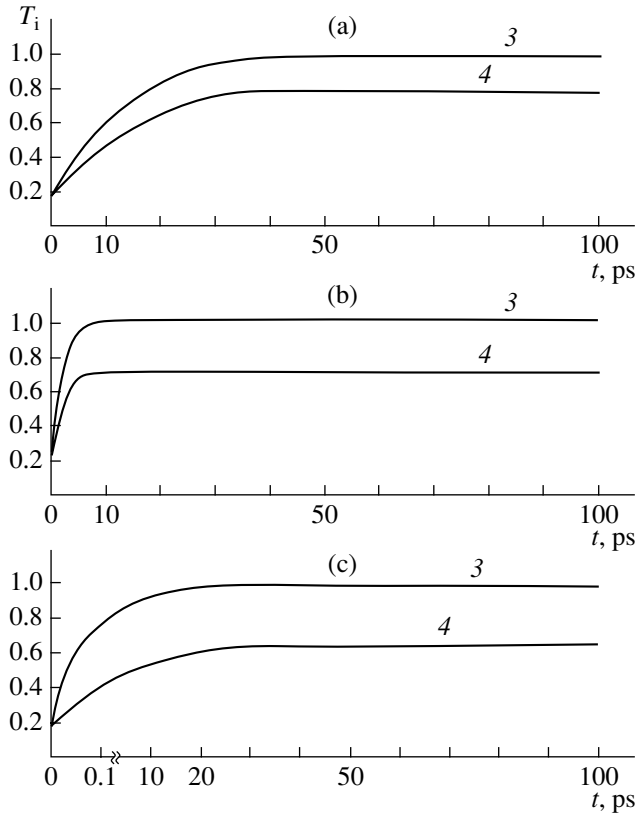


Fig. 3. Time dependence of the lattice temperature under the action of ultrashort laser pulses: (a)–(c) actions 2–4 (see table), respectively ($x = 0.45 \times 10^{-6}$ m). δ = (a) 22, (b) 49, and (c) 57%.

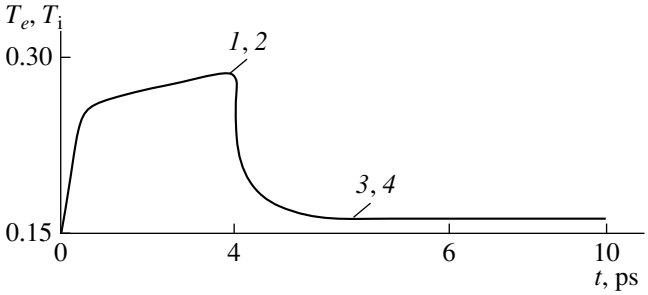


Fig. 4. Time dependences of the (*1*, *2*) electron temperature and (*3*, *4*) lattice temperature under the action of ultrashort laser pulses ($\tau_p = 4 \times 10^{-12}$ s and $I_0 = 0.8$ GW/cm^2).

when $\tau_p \le C_e/G$ (Figs. 2b, 3c). The reason is that, in this case, the major portion of the irradiation energy is converted to the emission current and is excluded from the energy balance in the system.

Thus, in the analysis of heating semiconductors exposed to ultrashort laser pulses, one can ignore thermionic emission until the sample begins to melt for the time $t = \tau_t$ if the exposure time exceeds the recombination time of the carriers, i.e., at $\tau_p > \tau_r$ (Fig. 1a).

In conclusion, it should be noted that when laser pulses are short (about several picoseconds) but have an intensity of less than $I_p = 35$ GW/cm^2, the effect of the emission is also insignificant ($\delta < 20\%$). For example, under the exposure conditions $\tau_p = 4 \times 10^{-12}$ s and $I_0 = 0.8$ GW/cm^2 [4], the electron and ion temperatures evaluated with and without the emission completely coincide ($\delta = 0\%$) (Fig. 4).

REFERENCES

1. V. É. Gusev, Kvantovaya Élektron. (Moscow) **11**, 2197 (1984).
2. S. M. Avanesyan and V. É. Gusev, Kvantovaya Élektron. (Moscow) **13**, 1241 (1986).
3. E. V. Gorbunov, N. A. Evtushenko, P. V. Lobzenko, and V. P. Sizov, Zh. Tekh. Fiz. **64** (4), 179 (1994) [Tech. Phys. **39**, 447 (1994)].
4. E. V. Gorbunov, Zh. Tekh. Fiz. **67** (5), 132 (1997) [Tech. Phys. **42**, 576 (1997)].
5. S. I. Anisimov, Ya. A. Imas, G. S. Romanov, and Yu. V. Khodyko, *Action of High-Power Radiation on Metals* (Nauka, Moscow, 1970).
6. N. B. Delone, *Interaction of Laser Emission with Materials* (Nauka, Moscow, 1989).
7. V. L. Bonch-Bruevich and S. G. Kalashnikov, *The Physics of Semiconductors: Textbook for Higher School* (Nauka, Moscow, 1990).
8. A. A. Samarskiĭ, *The Theory of Difference Schemes* (Nauka, Moscow, 1989).
9. R. E. Peierls, *Quantum Theory of Solids* (Clarendon Press, Oxford, 1955; Inostrannaya Literatura, Moscow, 1956).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 47, No. 1, 2002, pp. 74–79. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 76–82.

OPTICS, QUANTUM ELECTRONICS

Spontaneous Polarization and Nonlinear Susceptibility in Various Protonated $H_xLi_{1-x}NbO_3$ Phases

S. M. Kostritskii*, Yu. N. Korkishko, V. A. Fedorov**, M. V. Proyaeva**, and E. A. Baranov****

** Kemerovo State University, ul. Krasnaya 6, Kemerovo, 650043 Russia*

*** Moscow State Institute of Electronic Engineering (Technical University), Zelenograd, Moscow oblast, 103498 Russia*

e-mail: korkishk@chem.miee.ru

Received August 7, 2000; in final form, February 12, 2001

Abstract—A change in the spontaneous polarization and the nonlinear susceptibility in H : $LiNbO_3$ waveguides is calculated from experimental data for the shift of the fundamental absorption for variously polarized rays and from IR reflection data in the range of NbO_6 octahedron vibrations. It is shown that taking into account the lattice disorder provides good agreement between the theory and experiment. © *2002 MAIK "Nauka/Interperiodica".*

Because of the high nonlinear optical properties and high manufacturability, lithium niobate ($LiNbO_3$) is the most-used material in advanced nonlinear integrated optics. $LiNbO_3$-based optical waveguides are usually made by proton exchange (PE) [1].

While the PE technology is fairly simple, the waveguides produced have a variety of structures. It has been shown in our earlier works [2–6] that as many as seven $H_xLi_{1-x}NbO_3$ crystallographic phases may form in the waveguides, depending on proton exchange conditions and conditions of postexchange annealing. The dependences of the refractive index increment Δn_e for the extraordinary ray on the strain tensor normal component ε''_{33} on the waveguide surface have been constructed. Such dependences, which clearly illustrate all phase transitions (Fig. 1), have been called phase structural diagrams. It has been established that the β_4, β_3, β_2, β_1, and α phases of $H_xLi_{1-x}NbO_3$ can be obtained by either direct PE in the related solutions or annealing the previously prepared $H_xLi_{1-x}NbO_3$ phase with a higher proton concentration. Unlike these phases, the κ_1 and κ_2 phases cannot be produced by direct PE: they form as a result of the postexchange annealing of the β_i phases.

It is known that various $H_xLi_{1-x}NbO_3$ phases have different physical and physicochemical properties [6]. In [7], the quadratic coefficients of nonlinear distortion d were directly determined by measuring the second harmonic intensity when a focused laser beam (λ = 1.06 μm) scanned the polished end face of a waveguide. The phases β_4, β_3, and β_1 were found to have the extremely low nonlinearity: d_{33} (β_4, β_3, and β_1) $\le 0.1d_{33}$ ($LiNbO_3$). However, the β_2 phase demonstrates the much higher nonlinear distortion coefficients: $d_{33}(\beta_2) \approx 0.55d_{33}$ ($LiNbO_3$). This fact and also the rectangular refractive index profile in the β_i phases [6] make the β_2 phase promising for the production of second-harmonic-generating waveguides using the Vavilov–Cherenkov effect.

It has been shown [7] that postexchange annealing considerably increases the second harmonic intensity. However, radiation scattering appreciably degrades the reflected beam quality, making impossible the estimation of the nonlinear distortion coefficients in the κ_2, κ_1, and α phases obtained by postexchange annealing. In later investigations [8], it has been found that the nonlinear optical properties of α-H : $LiNbO_3$ waveguides produced by direct proton exchange at elevated temper-

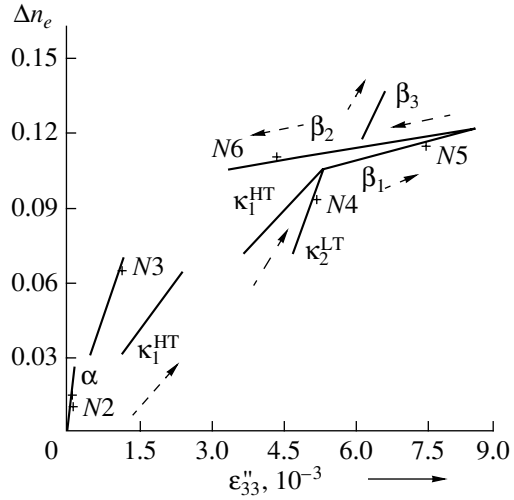


Fig. 1. Phase structural diagram for an *X*-cut $H_xLi_{1-x}NbO_3$ waveguide. Specimen positions are indicated.

1063-7842/02/4701-0074$22.00 © 2002 MAIK "Nauka/Interperiodica"

Table 1. Conditions for specimen preparation

Specimen no.	PE source	PE temperature, °C	PE time, h	Annealing temperature, °C	Annealing time, h	Surface phase
1	–	–	–	–	–	Pure $LiNbO_3$
2	Stearic acid + lithium stearate (0.75 wt %)	350	8	–	–	α
3	$KHSO_4$ solution in glycerol (4 g/l)	220	3	330	15	κ_1
4	$KHSO_4$ solution in glycerol (4 g/l)	220	3	330	5	κ_2
5	$KHSO_4$ solution in glycerol (1 g/l)	220	10	–	–	β_1
6	$KHSO_4$ solution in glycerol (4 g/l)	220	5	–	–	β_2

atures and those of initial lithium niobate [7] are nearly the same.

In this work, we try to disclose the reasons why the nonlinear optical properties of various $H_xLi_{1-x}NbO_3$ phases are so much different.

EXPERIMENT

PE waveguides were formed on *X*-cut lithium niobate substrates. The preparation conditions are listed in Table 1.

The effective refractive indices (ERIs) of the waveguide modes were measured at a wavelength of 633 nm with a prismatic coupling element. The RI distribution across the depth of the multimode single-phase waveguides was restored with the WKB method [9].

The strained state of the crystal structure was determined from X-ray rocking curves [10] obtained with a DRON-3 diffractometer. Then, using the measured values of the strains ε_{33} and the RI increments Δn_e, we determined the phase composition of the waveguides from the phase structural diagram [2, 3].

Transmission spectra in the visible and near-UV ranges were taken with a Specord 75 UV-VIS two-beam spectrophotometer. The radiation was incident normally to the specimens. Therefore, the transmission spectra taken from the specimens with the PE waveguides represented the superposition of the spectra from the crystal plate and from two thin (1.5–4.3 μm) near-surface $H_xLi_{1-x}NbO_3$ layers. Thus, the contribution of the surface layers can be separated out from the integral spectrum if the partial spectra of lithium niobate and the $H_xLi_{1-x}NbO_3$ phase considerably differ. The accuracy of discrimination is the highest when the transmission is measured near the fundamental absorption edge, where the optical density of the specimens is extremely high.

The IR spectra in the range of NbO_6 octahedron internal vibrations (300–1000 cm^{-1}) were taken with a Specord M80 spectrophotometer.

RESULTS AND DISCUSSION

The spectra taken from the $H_xLi_{1-x}NbO_3$-containing specimens and from the pure lithium niobate plates are depicted in Fig. 2. In the former case, the energy $E(k)$ of the fundamental absorption edge markedly decreases according to the phase formed. In the specimens containing any of the phases, this energy shift is different for the ordinary and extraordinary polarizations (Fig. 3). Moreover, the relationship between these shifts in orthogonal polarizations is not constant in con-

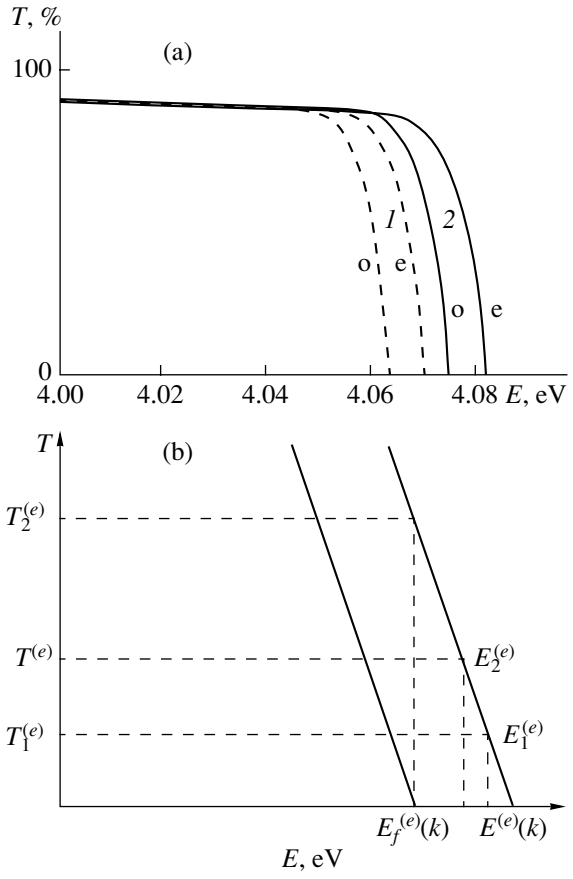


Fig. 2. Polarized radiation transmission spectra. (a) Optical transmission spectra for (*1*) specimen 4 with the κ_2 phase on the surface and (*2*) pure lithium niobate; (b) schematic diagram to determine the parameters used in the calculations with relationships (1) and (2).

Table 2. Oscillator strength shift in the various phases relative to pure lithium niobate

Specimen no.	Surface phase	Δn_e	$\Delta E_f^{(e)}$, eV	$\Delta E_f^{(0)}$, eV	$\Delta E_f^{(e)} - \Delta E_f^{(0)}$, eV
1	Pure $LiNbO_3$	0	0	0	0
2	α	0.01	0	0	0
3	κ_1	0.0645	−0.0449	−0.0362	−0.0087
4	κ_2	0.0935	−0.1098	−0.0884	−0.0214
5	β_1	0.1175	−0.0865	−0.0761	−0.0104
6	β_2	0.1100	−0.1187	−0.0924	−0.0263

trast to the predictions of the theory of polarization-induced energy band shift [11]. In the κ_i and β_i phases, this relationship is characterized by two constants (Fig. 3), which indicates that these phases have the fundamentally different microstructures.

As is known, the apparent fundamental absorption edge $E(k)$ depends on the reference beam attenuation during fixed-amplification photometric measurements. In our case, the reference beam was attenuated with a special diaphragm introduced into the reference channel. By measuring the beam attenuation in the absence of the specimen, we calibrated the transmission scale in arbitrary units. Since the photometric conditions remained the same for all of the specimens, we could estimate the strength E of the electronic oscillator responsible for the fundamental absorption in the specimen. The values of the strength for the ordinary and extraordinary polarizations diverge. The oscillator strengths $E_f^{(0)}$ and $E_f^{(e)}$ in the spectra of the

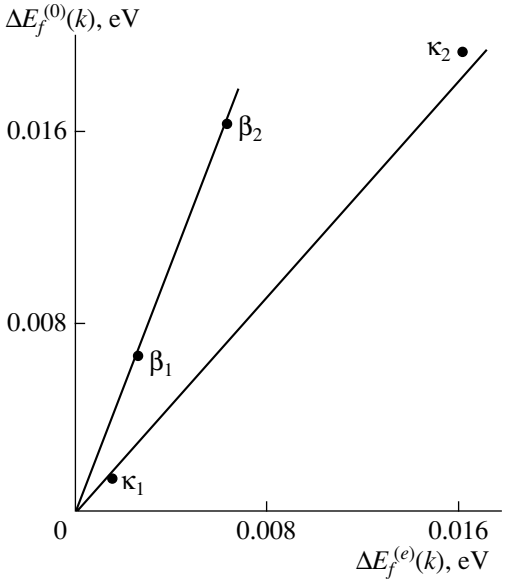


Fig. 3. Diagram showing the relationship between the shifts of the fundamental absorption edge for the ordinary and extraordinary polarizations, $\Delta E_f^{(0)}(k)$ and $\Delta E_f^{(e)}(k)$, with the different phases at the surface.

$H_xLi_{1-x}NbO_3$ phases were estimated for the Gaussian profile of the low-frequency fundamental absorption edge:

$$D^{(0,e)} = \alpha^{(0,e)}\exp\{-(E^{(0,e)} - E)^2/(\gamma^{(0,e)}/2)^2\}L,$$
$$D_f^{(0,e)} = \alpha^{(0,e)}\exp\{-(E_f^{(0,e)} - E)^2/(\gamma^{(0,e)}/2)^2\}2h. \quad (1a)$$

Taking the logarithm of the ratio $D_f^{(0,e)}/D^{(0,e)}$ for the values of E near the fundamental absorption edges, $E_f^{(0,e)}$, that are observed at the different polarizations, we obtain

$$E_f^{(0,e)} = E_f^{(0,e)}(k) + \{(E^{(0,e)} - E_f^{(0,e)}(k))^2 + (\gamma^{(0,e)}/2)^2\ln(2hD^{(0,e)}/LD_f^{(0,e)})\}^{1/2}. \quad (1b)$$

Here, $E_f^{(0,e)}(k)$ is the strength at the absorption edge in the spectrum of the specimen having two surface layers of the fth phase $2h$ thick, $\gamma^{(0,e)}$ is the attenuation coefficient for the electronic oscillator that is determined by extrapolating the experimental data, $D^{(0,e)}$ is the optical density in pure lithium niobate for the strength $E_f^{(0,e)}(k)$, $D_f^{(0,e)} = D_f^{(0,e)}(k) - D^{(0,e)}$ is the optical density of the specimen for the two layers of the fth phase under study, $D_f^{(0,e)}(k)$ is the optical density of the specimen for the strength $E_f^{(0,e)}(k)$, L is the thickness of the crystal plate (1 mm in our case), and $\alpha^{(0,e)}$ is the absorption coefficient in the maximum of the fundamental absorption [11] (that is, at $E = E^{(0,e)}$ or $E_f^{(0,e)}$).

The values of $\gamma^{(0,e)}$ were found from the spectra measured for the Gaussian profile of the fundamental absorption band:

$$\gamma^{(0,e)} = 2((E_1^{(0,e)} - E_2^{(0,e)}) \times (2E^{(0,e)} - E_1^{(0,e)} - E_2^{(0,e)})/\ln(D_1^{(0,e)}/D_2^{(0,e)}))^{1/2}, \quad (2)$$

where $D_1^{(0,e)}$ and $D_2^{(0,e)}$ are the optical densities of pure lithium niobate for two arbitrary values of the energy: $E_1^{(0,e)} > E_2^{(0,e)}$.

From the spectra taken, it follows that $\gamma^{(0)} = 0.547$ eV and $\gamma^{(e)} = 0.702$ eV. In the calculation, the previously determined [11] and widely used values of the oscillator strength were employed: $E^{(0)} = 6.74$ eV and $E^{(e)} = 7.08$ eV. Equation (1) allows us to estimate the oscillator strength shift $\Delta E_f^{(0,e)} = E_f^{(0,e)} - E^{(0,e)}$ in the phase studied compared to pure lithium niobate (Table 2).

It should be noted that the oscillator strength shift ΔE_f for the extraordinary polarization was larger than for the ordinary one in all of the phases. At the same time, the shift of the fundamental absorption edge $\Delta E_f(k)$ is larger for the ordinary polarization. This is because the oscillator strength for the ordinary polarization is lower; hence, a smaller energy shift changes the optical density to a greater extent than in the case of the extraordinary polarization. The relationship between the oscillation strength shifts for the two polarizations is consistent with the theory [11]. After normalization to the optical path length in the associated phase, i.e., to $2h$, the values of $\Delta E_f^{(0,e)}$ in the κ_2 phase turn out to be larger than in the β_2 phase, whereas for the shifts of the absorption edge $\Delta E_f(k)$, the inverse relationship holds (cf. Tables 2, 3).

According to the theory of energy band shifts in ferroelectrics with oxygen-filled octahedrons [10], the shifts in the oscillator strength imply the significant decrease in the spontaneous polarization P in the $H_xLi_{1-x}NbO_3$ phases in comparison with that in pure $LiNbO_3$. It is known [11, 12] that a change in the spontaneous polarization P_0 by a value of ΔP causes an oscillator strength shift $\Delta E^{(0,e)}$ that is defined by the associated component of the effective polarization potential tensor β_{ij}:

$$\Delta E^{(e)} = \beta_{11}\Delta P P_0, \quad \Delta E^{(0)} = \beta_{12}\Delta P P_0, \qquad \Delta E^{(e)} - \Delta E^{(0)} = \beta_{44}\Delta P P_0. \tag{3}$$

Thus, using (3), one can determine the spontaneous polarization P_f ($P_f = P_0 + \Delta P$) in the phase studied from the oscillator strength shift. For example, from the empirical values of the component β_{11} [12], we determined the spontaneous polarization P_f in the related phase (Table 3).

The relative nonlinear susceptibilities d_f in the various phases were estimated by the formula [11]

$$d_f/d_0 = (\beta_{11,f} P_f (E^{(e)})^3)/(\beta_{11} P_0 (E_f^{(e)})^3), \tag{4}$$

where d_0 is the associated value in pure lithium niobate.

Table 3 lists the nonlinear susceptibilities for the $H_xLi_{1-x}NbO_3$ phases that were calculated under the assumption that the effective polarization potential remains constant ($\beta_{11,f} = \beta_{11}$) Note that our data for the changes in the ratio $(\Delta E_f^{(e)} - \Delta E_f^{(0)})/\Delta E_f^{(e)}$ and also the data for the shifts of the fundamental absorption edge (Fig. 3) indicate that the components β_{12} and β_{44} of the effective polarization potential change. From these data, we cannot accurately estimate this change. However, it may be safely suggested that it is small. Therefore, we determined only one coefficient of nonlinear susceptibility, using β_{11} from [12] and experimental values of $E_f^{(e)}$ for various $H_xLi_{1-x}NbO_3$ phases.

Table 3. Relative spontaneous polarizations and nonlinear susceptibility coefficients calculated with the transmission spectrum parameters determined from the local response of the crystal

Phase	P_f/P_0	d_f/d_0
α	1.0	1.0
κ_1	0.907	0.924
κ_2	0.772	0.809
β_1	0.820	0.851
β_2	0.754	0.793

It should be noted that the above values of P_f/P_0 and d_f/d_0 far exceed our data obtained from IR reflection spectra for various phases of lithium niobate [13] and also by directly measuring the second harmonic intensity [7]. Recall, however, that the above calculations allow the determination of a local response, while direct measurements reflect the superposition of the contributions from many local oscillators. Obviously, the lattice disorder in the protonated phases would reduce the coherency of local oscillators; consequently, the integral response that is determined by directly measuring P_f and d_f may diminish and even disappear. In fact, in the proton-substituted phases, especially in the κ_2 and β_1, the lattice disorder is significant, as follows from our data for IR reflection spectra. From these spectra, the optical phonon attenuation in the various phases was determined.

For all of the proton-substituted phases, as well as for pure lithium niobate, the IR reflection spectra exhibit the noticeable attenuation coefficient Γ; namely, the reflection coefficient R peaks at frequencies much higher than the transverse optical (TO) phonon frequencies ω(TO), remaining much less than unity (Fig. 4). In the 580–890 cm^{-1} range of the IR reflection spectra, which corresponds to the phonon mode of NbO_6 octahedron internal vibrations, $R_{max} = 0.938$, 0.90, 0.85, 0.74, and 0.90 for the α, κ_1, κ_2, β_1, and β_2 phases, respectively. These values were obtained on the X-cut specimens. Thus, the dependence of R_{max} on the proton concentration x in $H_xLi_{1-x}NbO_3$ waveguides is nonmonotonic (Fig. 5). Note that the lattice strain determined by X-ray diffraction analysis in the various phases depends on the proton concentration. The associated relationship has been found earlier [13]. It is also known that the greater Γ, the smaller R_{max} [14]. Unfortunately, the analytical relationship between these

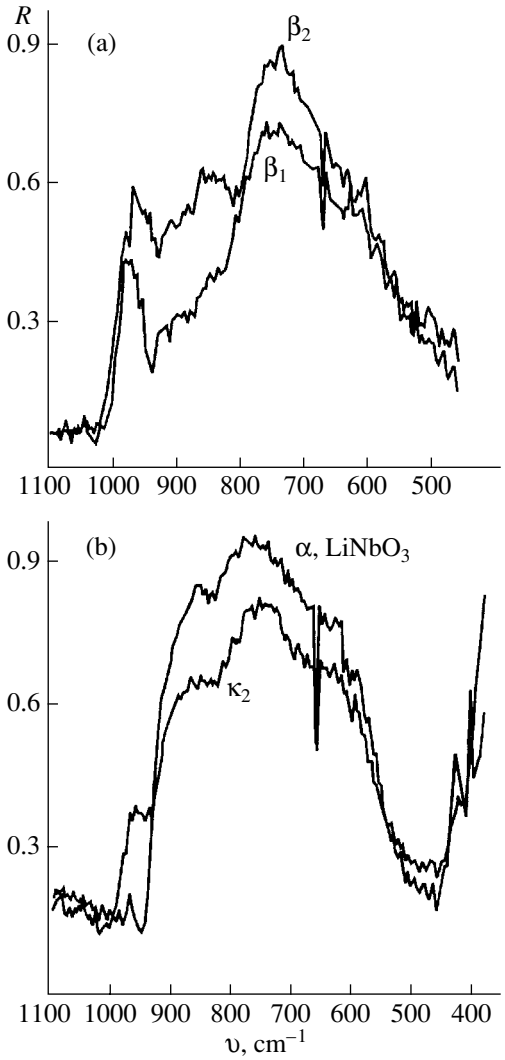


Fig. 4. IR reflection spectra for $H_xLi_{1-x}NbO_3$ waveguides with (a) high and (b) low proton concentration x.

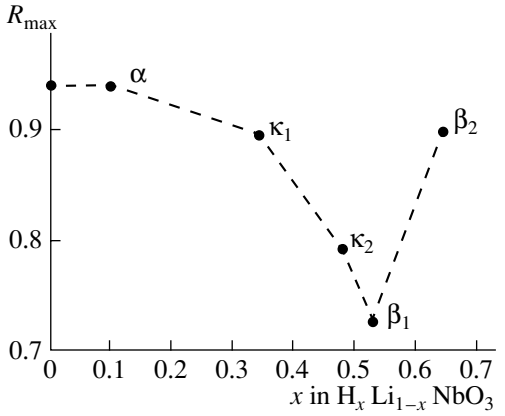


Fig. 5. Maximal reflection coefficient R_{max} in the range 580–890 cm^{-1} vs. relative proton concentration x in the $H_xLi_{1-x}NbO_3$ phases. The given band corresponds to the optical phonon mode for the internal oscillations of the NbO_6 octahedron.

parameters includes a number of undeterminable quantities. However, there exists a simple analytical dependence between the reflection coefficient $R(\omega(TO))$ at the transverse phonon frequency and the attenuation coefficient Γ of this phonon [14]. With this dependence, we obtained the following formula for Γ in the protonated phase:

$$\Gamma_f = \Gamma_0(\{1 - R_f(\omega_f(TO))\}^2 \times \omega_0(TO))/(\{1 - R_0(\omega_0(TO))\}^2\omega_f(TO)), \quad (5)$$

where the subscripts f and 0 refer to the protonated phase and pure lithium niobate, respectively.

The value of $\Gamma_0 = 15$ cm^{-1} was taken from the literature [15], and the other parameters were determined from the IR reflection spectra. For the α, κ_1, κ_2, β_1, and β_2 phases, we obtained $R_f(\omega_f(TO)) = 0.79$, 0.76, 0.65, 0.56, and 0.68, respectively; for pure congruently melting lithium niobate, $R_0(\omega_0(TO))$ was found to be 0.79. Substituting the experimental data into (5) yields $\Gamma_f =$ 15, 19, 38, 59.5, and 28.4 cm^{-1} for the α, κ_1, κ_2, β_1, and β_2 phases, respectively.

It is known [16] that the value of Γ for optical phonons increases with the lattice disorder; in other words, Γ is inversely proportional to the mean correlation radius of the structure R_c: $\Gamma \sim a/(\pi R_c)$, where π/a is the radius of the first Brillouin zone and $1/R_c$ is the correlation vector in the reciprocal space. Bearing in mind that the temperature-dependent optical phonon attenuation component due to oscillation anharmonicity [17, 18] also grows with decreasing long-range correlation radius (i.e., with decreasing R_c) [16], we come to

$$(R_c)_f/(R_c)_0 \sim \Gamma_0/\Gamma_f. \quad (6)$$

Note that the value of Γ in our congruently melting crystals is relatively large because of the high concentration (5.7–6.0 mol %) of intrinsic defects, such as lithium vacancies and Nb_{Li} antisite defects [17, 18]; in other words, the value of R_c is to a large extent limited by the violation of long-range order, since the mean interdefect spacing is small.

Assuming that the coherence distance L_c within which the phase matching of secondary waves in a nonlinear crystal is retained equals R_c, one can estimate the additional factor F that is responsible for the decrease in the intensity of the integral response in experiments on second harmonic generation (SHG) when the coherence of superposing local responses fails. According to [19], $F \sim 1/(L_c)^2$; therefore, only the relative value of F can directly be found from IR spectroscopy data for the various phases. The value of F_f/F_0 can be estimated from the relationship

$$F_f/F_0 = (\Gamma_f/\Gamma_0)^2. \quad (7)$$

Substituting Γ_f and Γ_0 found from the IR reflection spectra into (7) yields $F_f/F_0 = 1.0$, 1.6, 6.4, 15.7, and

Table 4. Nonlinear susceptibility coefficients for the nonlinear integral response that are calculated with regard for the data on lattice disorder in the various $H_xLi_{1-x}NbO_3$ phases (values obtained from the second harmonic intensity measurements are given for comparison)

Phase	d_f/d_0	
	experiment	calculation
α	1.0	1.0
κ_1	–	0.73
κ_2	–	0.28
β_1	0.1	0.12
β_2	0.55	0.5

3.58 for the α, κ_1, κ_2, β_1, and β_2 phases, respectively. Thus, the effective nonlinear susceptibility $(d_f)_{\text{eff}}$ measured directly in the SHG experiments may be much lower than spectroscopically determined d_f, which characterizes the nonlinear optical response. These values relate as

$$(d_f)_{\text{eff}}/(d_f) = 1/F_f. \quad (8)$$

To find the moduli of the factor F_f, we used calibration by $(d_f)_{\text{eff}}$ for the β_2 phase. This parameter was determined from the direct measurements of the second harmonic intensity in proton-substituted waveguides [4]. The results were the following: (1) When for the local response in the β_2 phase, we used d_f obtained from the IR spectroscopy data [12], F_f = 1, 1.23, 2.50, 6.12, and 1.40 for the α, κ_1, κ_2, β_1, and β_2 phases, respectively, and (2) with d_f obtained from the shift of the fundamental absorption edge, F_f = 1, 1.27, 2.88, 7.07, and 1.59 for the same phases.

Table 4 lists the experimental nonlinear susceptibilities and those calculated with regard for the lattice disorder. It is clearly seen that, with the lattice disorder taken into account in the calculation of the nonlinear optical response, the agreement with the experimental data for the nonlinear optical properties of the $H_xLi_{1-x}NbO_3$ waveguides is fairly good. Thus, we suggested the efficient technique for calculating the spontaneous polarization and the nonlinear susceptibility coefficients in $LiNbO_3$-based PE waveguides. Our technique relies upon experimental data for the shift of the fundamental absorption edge and IR reflection spectroscopy. It can also be applied for analyzing other ferroelectric crystals.

CONCLUSION

It has been found that lattice disorder is the basic factor among those responsible for the discrepancy between the integral and local nonlinear optical responses in $LiNbO_3$-based PE waveguides. Two reasons govern the dependence of the nonlinear optical properties of these waveguides on their phase composition: (1) a change in the strength of local electronic oscillators and (2) the incoherence of the contributions from the local oscillators to the integral nonlinear optical response because of the lattice disorder in the $H_xLi_{1-x}NbO_3$ phases.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 96-15-96978, 99-15-96110, and 96-15-98220), the Program “Russian Universities: Basic Research” (grant no. 248-1998), and the Federal Program “Integration” (project no. A-102).

REFERENCES

1. J. L. Jackel, C. E. Rice, and J. J. Veselka, Appl. Phys. Lett. **41**, 607 (1982).
2. Yu. N. Korkishko and V. A. Fedorov, Kristallografiya **44**, 238 (1999) [Crystallogr. Rep. **44**, 237 (1999)].
3. Yu. N. Korkishko and V. A. Fedorov, IEEE J. Sel. Top. Quantum Electron. **2**, 187 (1996).
4. Yu. N. Korkishko and V. A. Fedorov, Zh. Tekh. Fiz. **69** (3), 47 (1999) [Tech. Phys. **44**, 307 (1999)].
5. Yu. N. Korkishko, V. A. Fedorov, M. P. De Micheli, *et al.*, Appl. Opt. **35**, 7056 (1996).
6. Yu. N. Korkishko and V. A. Fedorov, Cambridge Intern. Sci. Publ. (1999), p. 97.
7. Yu. N. Korkishko, V. A. Fedorov, and F. Laurell, IEEE J. Sel. Top. Quantum Electron. **5**, 101 (2000).
8. Yu. N. Korkishko, V. A. Fedorov, and O. Y. Feoktistova, J. Lightwave Technol. **18**, 562 (2000).
9. J. M. White and P. F. Heidrich, Appl. Opt. **15**, 151 (1976).
10. V. A. Gan’shin, Yu. N. Korkishko, and V. A. Fedorov, Kristallografiya **40**, 341 (1995) [Crystallogr. Rep. **40**, 308 (1995)].
11. Yu. S. Kuz’minov, *Ferroelectric Crystals for Laser Emission Control* (Nauka, Moscow, 1982).
12. V. N. Fradkin, K. D. Kochev, Yu. S. Kusminov, *et al.*, Phys. Status Solidi **33**, K137 (1976).
13. Yu. N. Korkishko, V. A. Fedorov, and S. M. Kostriskii, J. Appl. Phys. **84**, 2411 (1998).
14. H. Poulet and J.-P. Mathieu, *Spectres de vibration et symétrie des cristaux* (Gordon and Breach, Paris, 1970; Mir, Moscow, 1973).
15. I. P. Kaminow and W. D. Johnston, Phys. Rev. **160**, 519 (1967).
16. H. Vogt, J. Phys.: Condens. Matter **3**, 3697 (1991).
17. S. Kojima, Jpn. J. Appl. Phys. **32**, 4373 (1993).
18. U. T. Schwarz and M. Maier, Phys. Rev. B **55**, 11 041 (1997).
19. R. C. Miler, J. Phys. Soc. Jpn. **28**, 15 (1970).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 80–87. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 83–91.

RADIOPHYSICS

The Generation of Superradiance Pulses by High-Current Subnanosecond Electron Bunches Moving in a Periodic Slow-Wave System: Theory and Experiment

N. S. Ginzburg*, I. V. Zotova*, Yu. V. Novozhilova*, A. S. Sergeev*, V. G. Shpak, S. A. Shunaĭlov**, M. R. Ul'maskulov**, and M. I. Yalandin****

** Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603600 Russia*

*** Institute of Electrophysics, Ural Division, Russian Academy of Sciences, Komsomol'skaya ul. 24, Yekaterinburg, 620016 Russia*

e-mail: river@appl.sci-nnov.ru

Received April 4, 2001

Abstract—Cherenkov superradiance observed when an electron bunch rectilinearly moves through a slow-wave periodic system is studied theoretically and experimentally. The simulation based on averaged equations and the direct numerical simulation using the PIC-code KARAT show that the peak power of the microwave pulses varies as the total number of the particles in a bunch squared. This finding is confirmed experimentally. Ultrashort (300 ps wide) high-power (up to 140 MW) pulses are generated at a frequency of 39 GHz. As an electron source, the high-current subnanosecond RADAN-303 accelerator is used. It injects 0.5- to 1.5-ns-wide electron bunches of current up to 2 kA and energy 200–300 keV. The simulation suggests that the power of the electromagnetic pulses can be increased further (up to 300 or 400 MW) by optimizing the accelerating voltage pulse shape. © *2002 MAIK "Nauka/Interperiodica".*

The analysis of the dynamics of backward-wave oscillators [1, 2] has demonstrated that the stationary oscillation mode changes to the mode of self-modulation as the injection current grows. It has been noted that, early at the transition stage, a high-power electromagnetic pulse is generated. Its amplitude far exceeds the oscillation amplitude predicted by the stationary theory for equal injection currents [3]. It has been shown that this pulse is a result of the interaction between different parts of the electron flow. This interaction is due to a wave slip relative to the electrons, since the group velocity of the wave differs from the translational velocity of the particles. The idea to employ this mechanism for generating high-power ultrashort microwave pulses has been put forward [4]. In this case, the electron beam duration must be limited by a time interval on the order of the relaxation time

$$T_c = l\left(\frac{1}{V_0} + \frac{1}{V_{gr}}\right). \quad (1)$$

Here, l is the characteristic length of the space of interaction, $V_0 = \beta_0 c$ is the translational velocity of the particles, and $V_{gr} = \beta_{gr} c$ is the group velocity of the wave.

It should be emphasized that such single pulses can be viewed as superradiance (SR) pulses, because some of their properties are typical of SR pulses (cf. [5–8]). One of them is the quadratic dependence of the peak power on the total number of particles in a bunch. Such a dependence means that the amplitudes of the fields radiated by most electrons in the bunch may add coherently at some time instant as a result of particle grouping. As the number of particles in the bunch grows, the pulse generated shortens.

The pioneering experiments on SR pulse generation at frequencies near 40 GHz using this mechanism were carried out in 1997 [9, 10]. In them, the electron bunches were focused with the relatively low (<2 T) magnetic field of an impulse solenoid. This field strength was lower than that corresponding to the cyclotron absorption band. At the same time, early investigations into long-pulse relativistic backward-wave tubes showed that these sources have two ranges of efficient oscillation separated by the cyclotron absorption band [11, 12]. With this in mind, one might expect that, in the presence of strong magnetic fields (exceeding the resonance value), the peak oscillation power will also be several times higher than that obtained in the early experiments when short electron pulses are generated. In this case, the power increases primarily because the quality of the bunch formed by an explosive emission cathode is improved. Additionally, electron focusing by a strong magnetic field helps to bring the slow-wave system closer to the bunch, increasing the coupling resistance and associated increments. In the next series of experiments, a supercon-

1063-7842/02/4701-0080$22.00

ducting solenoid was used to create a permanent magnetic field as high as 10 T. The application of the high magnetic field allowed us to increase the peak radiated power to 60 MW and bring the pulse repetition rate to 25 Hz [13, 14]. At the last stage of those investigations, we succeeded in rising the energy of the beam and the injection current to 290 keV and 2 kA, respectively, by modifying the accelerating pulse former of the subnanosecond electron injector. The impulse solenoid generated a guiding magnetic field of up to 5 T. As a result, the peak power of SR pulses grew to 140 MW.

In this work, we experimentally and theoretically study the mechanism of SR pulse generation when a rectilinearly moving electron bunch interacts with a slow spatial harmonic of the backward wave propagating in the slow-wave system.

SR PULSE GENERATION DUE TO BACKWARD WAVE–ELECTRON INTERACTION: THEORETICAL ANALYSIS

The electrons interact with a concurrent spatial harmonic of the backward wave if the synchronism condition (Fig. 1)

$$\omega = (-h + \bar{h})V_0 \qquad (2)$$

holds. Here, ω and h are the frequency of the wave and the longitudinal wave number, respectively, and $\bar{h} = 2\pi/d$ (d is the corrugation period).

Then, the longitudinal electric field can be represented as

$$E_z = \mathrm{Re}[E_z^s(\mathbf{r}_\perp)A(z, t)\exp(i\omega(t - z/V_0))], \qquad (3)$$

where $E_z^s(\mathbf{r}_\perp)$ is the transverse field distribution that coincides with a mode of the corrugated waveguide and $A(z, t)$ is the slowly varying amplitude of the wave.

Electron–wave interaction is then described by the set of equations that includes the equation for the excited wave amplitude,

$$\frac{\partial a}{\partial \tau} - \frac{\partial a}{\partial \zeta} = -f(\tau)\frac{G}{\pi}\int_0^{2\pi} e^{-i\Theta}d\Theta_0 \qquad (4)$$

and the motion equations for relativistic electrons:

$$\frac{\partial \Theta}{\partial \zeta} = \frac{1}{\sqrt{1-\gamma^{-2}}} - \frac{1}{\sqrt{1-\gamma_0^{-2}}}, \qquad (5)$$

$$\frac{\partial \gamma}{\partial \zeta} = \mathrm{Re}(a(\zeta, \tau)e^{i\Theta}). \qquad (6)$$

Here, we introduced the following dimensionless parameters and variables: $a = eE_z^s(r_b)A/(mc\omega)$; $\tau = \omega(t - z/V_0)(1/\beta_0 + 1/\beta_{dr})^{-1}$ and $\zeta = \omega z/c$ are time and coordinate, respectively; $\Theta = \omega t - h_s z$ is the phase of an electron relative to the synchronous harmonic of the field; $h_s = \bar{h} - h$; $G = eI|Z|/(2\beta_0^2 mc^2)$; Z is the coupling resistance for this harmonic [15]; I is the beam current; $\gamma = (1 - \beta_0^2)^{-1/2}$ is the relativistic mass factor of the electrons; and ω is the frequency of strict synchronism.

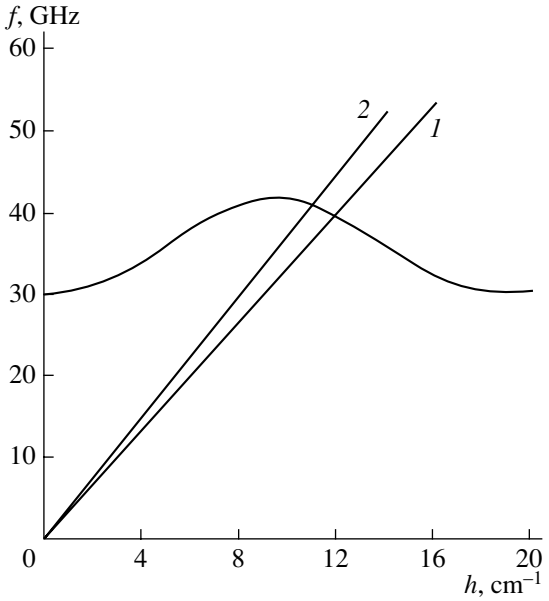


Fig. 1. Dispersion diagram for Cherenkov synchronism with the backward wave in a periodic slow-wave system made as a corrugated waveguide. The corrugation period and depth are 3.5 and 0.75 mm, respectively, and the mean radius of the waveguide is 3.75 mm. The electron energy is (*1*) 200 and (*2*) 300 keV.

The function $f(\tau)$ stands for the density distribution in the bunch. We assume that $f(\tau) = 1$ for $\tau \in [0, T_e]$; that is, the electron density is uniformly distributed in the bunch of dimensionless duration

$$T_e = \omega t_{\mathrm{e.pulse}}(1/\beta_0 + 1/\beta_{\mathrm{gr}})^{-1}. \qquad (7)$$

The boundary conditions to Eqs. (4)–(6) have the form

$$\Theta|_{\zeta=0} = \Theta_0 + r\cos\Theta_0, \quad \Theta_0 \in [0, 2\pi], \quad \gamma|_{\zeta=0} = \gamma_0,$$
$$a|_{\zeta=L} = 0, \quad a|_{\tau=0} = a_0(\zeta), \qquad (8)$$

where $L = \omega l/c$ is the dimensionless length of the space of interaction and $r \ll 1$ is the parameter that includes initial fluctuations of the electron bunch density.

Figure 2 shows the results of the numerical simulation of Eqs. (4)–(6) for the electron pulse duration $T_e = 25$, slow-wave system length $L = 50$, and various values of the parameter G characterizing the bunch current. Note first of all that the radiation is single-pulse in this range of the parameters (Fig. 2a). From Fig. 2b, it follows that the peak amplitude of SRPs linearly varies with current at small G or, in other words, that the peak power of the radiation quadratically depends on the total number of the particles in the bunch. This means

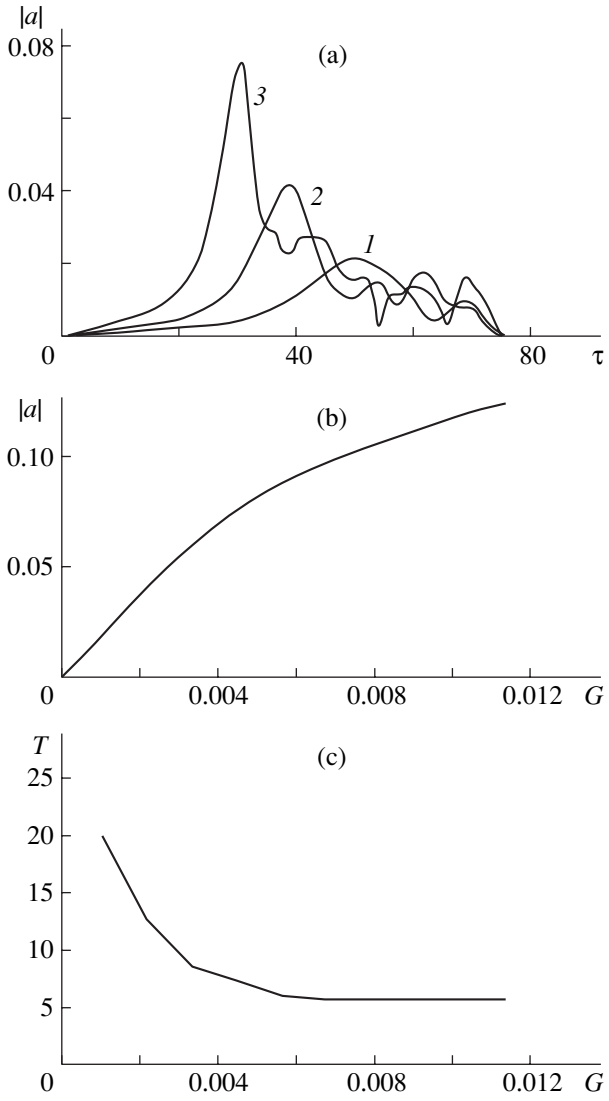


Fig. 2. Dependence of the (a) SR pulse shape, (b) peak amplitude, and (c) width on the electron bunch current. G = (*1*) 0.001, (*2*) 0.002, and (*3*) 0.004. The normalized duration of the electron bunch is T_e = 25.

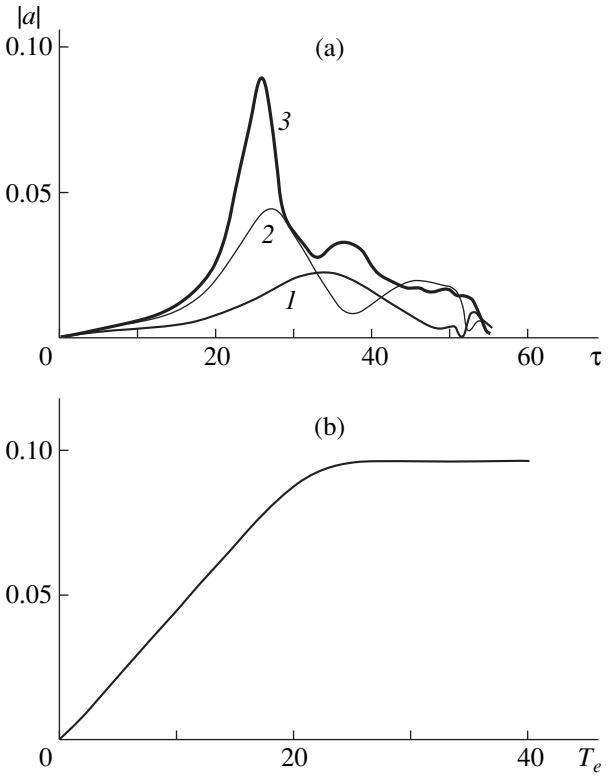


Fig. 3. Dependence of the (a) SR pulse shape and (b) peak amplitude on the electron bunch duration. G = 0.008. T_e = (*1*) 5, (*2*) 10, and (*3*) 20.

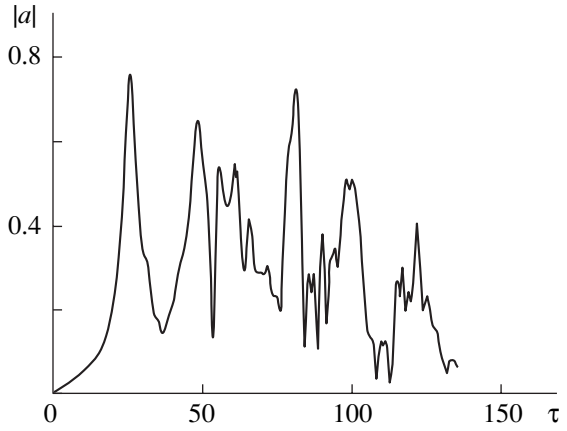


Fig. 4. Multipulse mode of radiation when the duration of the electron bunch is much larger than the correlation length. T_e = 100 and L = 50.

that all electrons of the bunch radiate coherently. As the number of the particles in the bunch grows, not only the peak power increases but the pulses shorten as well (Fig. 2c). Such a behavior is typical of SR pulses [5].

The dependence of the SR pulse amplitude on the electron bunch duration at the fixed G = 0.008 (i.e., at the fixed current) (Fig. 3) is, like the current dependence, also linear for as long as the bunch duration is small. At $T_e > 25$, the peak power saturates, because the electron pulse becomes too long and can no longer provide the coherency of the radiation from the entire bunch. The radiation becomes multipulse (Fig. 4). As the electron bunch widens, we virtually observe the transition to the self-modulation mode of oscillation, which has been studied upon feeding backward-wave tubes by quasi-stationary electron beams [1, 2].

The absolute values of the peak powers and SR pulse durations were estimated under conditions similar to those of the above experiments. The length of the slow-wave system was 6 cm; the mean radius of the waveguide, 3.75 mm; the corrugation period, 3.5 mm; and the depth of corrugation, 0.75 mm. Based on the

experiments performed, we considered two cases. In the former case, the electron energy and current were equal to 200 keV and ≈1 kA, respectively; in the latter, 300 keV and ≈2 kA. The frequency of synchronism was 40 and 42 GHz (Fig. 1). The group velocity of the wave at the points of synchronism was estimated from the dispersion diagram and was found to be about 0.3*c*. Then, for an electron pulse duration of 500 ps, the normalized duration T_e is approximately 25. With these parameters, the normalized length of the space of interaction was $L \approx 50$. The peak radiation density can be estimated by the formula

$$P = \frac{m^2c^5}{e^2}\frac{eI_0}{mc^3}\frac{|a|^2}{4G}. \quad (9)$$

Figure 5 shows the peak radiation power in arbitrary units vs. the coupling resistance *Z*. For a spacing between the electrons and the slow-down system of 0.25 mm, the coupling resistance evaluated by the method of integral equations [15] was found to be $Z \approx 1.3\ \Omega$. As follows from Fig. 5, the peak power in this case reaches ≈100 MW at an electron energy of 200 keV and a current of 1 kA or ≈250 MW for the particle energy 300 keV and the current 2 kA. In both cases, the SR pulse duration was about 300 ps, which corresponds roughly to 15 periods of high-frequency oscillations.

Thus, our analysis demonstrates that the superradiance of an electron bunch moving in a periodic slow-wave system may show up as the generation of high-power nanosecond electromagnetic pulses in the millimeter range.

EXPERIMENTAL SETUP

The nanosecond high-voltage RADAN-303 generator with the subnanosecond pulse sharpening device was used as the basic unit of the electron bunch injector (0.5–1.5 ns, 1–2 kA, and 200–300 keV) [16, 17]. The tubular (mean radius 2.75 mm, wall thickness 0.4 mm) bunches were generated by the coaxial magnetically insulated explosive emission diode. The subnanosecond signals from the voltage, current, and microwave power transducers were recorded by a Tektronix 7250 oscilloscope with a record band of 7 GHz.

The high-current electron bunches were transported through the 6-cm-long space of interaction either by the permanent magnetic field (up to 10 T) of the superconducting solenoid or by the field (up to 5 T) of the impulse solenoid. The working space was a part of the corrugated waveguide with the parameters given above. The depth of corrugation was smoothly decreased in the last three periods toward the collector end in order to provide matching with the output electrodynamic path. Similarly, the depth of corrugation was decreased in the first four periods toward the cathode end to provide matching with the evanescent taper. After reflecting from the evanescent taper, the electromagnetic pulse generated was de-excited in the direction of the translational motion of the electrons.

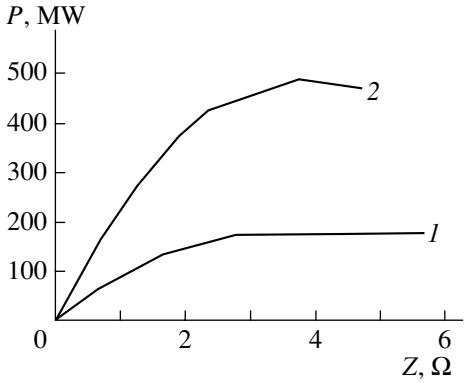


Fig. 5. Absolute value of the SR pulse peak power vs. coupling resistance. (*1*) γ = 1.4, I_0 = 1 kA and (*2*) γ = 1.6, I_0 = 2 kA.

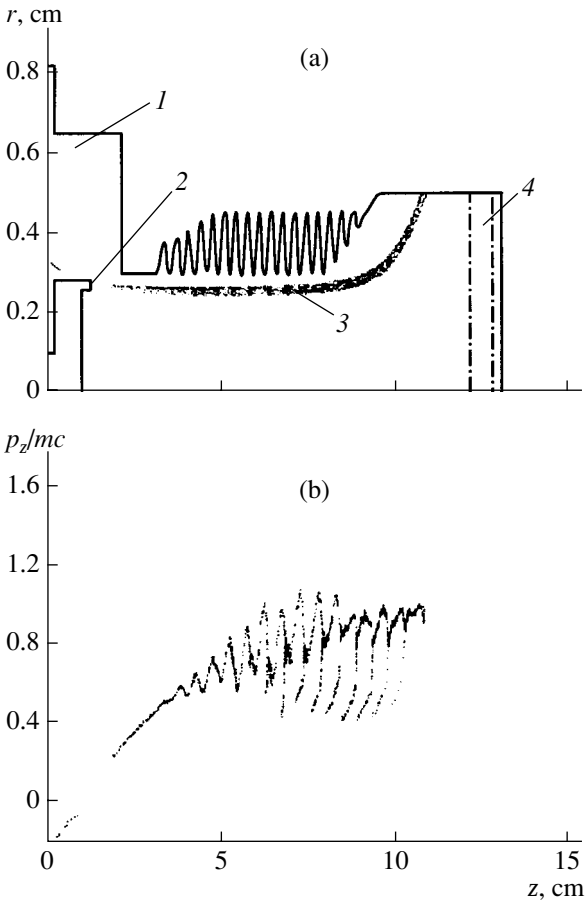


Fig. 6. (a) Geometry of the space of interaction used in KARAT-code-based simulation: *1*, feeding coaxial line; *2*, injecting area; *3*, arrangement of the particles 1 ns after an accelerating voltage pulse has been applied to the coaxial line; and *4*, absorber. (b) Arrangement of the particles on the phase plane (p_z, *z*) for *t* = 1 ns.

Recall that the focusing magnetic field strengths used in the experiments (5 T or higher) exceeded those

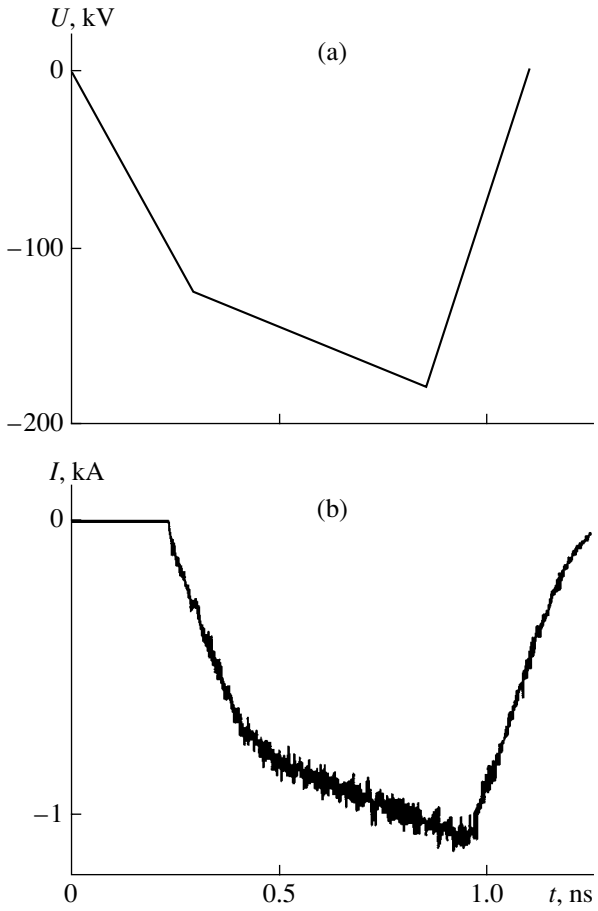


Fig. 7. Shape of the (a) accelerating voltage pulse and (b) self-consistent current (simulation).

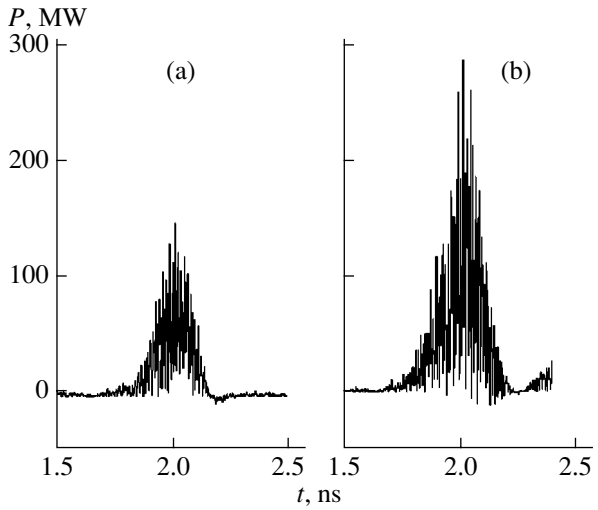


Fig. 8. SR pulses for different accelerating voltage amplitudes (simulation). U_{max} = (a) 180 and (b) 290 kV.

corresponding to the absorption band. As is known, the absorption band arises when the cyclotron resonance condition for the fundamental harmonic of a wave propagating in a periodic system is met. In addition, the transport of the bunch in the high magnetic field allowed us to shrink the gap between the slow-wave system and the particles nearly to one tenth of a millimeter and thereby increase the coupling resistance and the increments.

NUMERICAL SIMULATION OF ELECTRON BUNCH SUPERRADIANCE USING THE KARAT CODE

The superradiance of an electron bunch moving in a periodic slow-wave system was additionally simulated with the PIC-code KARAT, which makes it possible to directly (without averaging) integrate the Maxwell equations jointly with the motion equations for the electrons. The geometry of the system (Fig. 6a) was taken close to actual experimental conditions. The simulation included the formation of an electron pulse in a magnetically insulated coaxial diode. The self-consistent model of the diode was applied. The emitter was specified in the form of a half-toroid on the end face of the hollow cathode, which has a radial thickness of ≈0.2 mm. The grid pitch (≈0.05 mm) covered three or four annular zones on the emitting surface. If a 180-kV 1-ns-wide voltage pulse (Fig. 7a) is applied to the feeding coaxial line (the geometry shown in Fig. 6a), the electron energy at the entrance to the working space is ≈200 keV at a current pulse amplitude of ≈1.1 kA (Fig. 7b). These values agree well with experimental data. During simulation, the magnetic field strength was taken to be 5 T.

For the corrugation parameters given above, the transit of the bunch through the space of interaction causes electron modulation in terms of density (Fig. 6a) and longitudinal momentum (Fig. 6b). The result is the de-excitation of a short electromagnetic pulse presented in Fig. 8a. The half-height width of the SR pulse is no more than 300 ps for a peak power of up to 75 MW (note that the power averaged over the short period is half as large as the instantaneous power shown in Fig. 8). The output radiation represents the TM_{01} wave with a center frequency of 39 GHz (Fig. 9).

Figure 10 plots the peak power against the voltage pulse duration. Since the peak value of the electron current remains practically constant and the width of the electron pulse depends on that of the accelerating voltage pulse, this dependence, in essence, demonstrates how the peak power varies with the total charge of the electron bunch. The curve in Fig. 10 is nearly quadratic for as long as the electron pulse width is smaller than the correlation time (for the parameter values considered, T_c is estimated at ~1 ns from (1)).

Note that, if the electron pulse is short, the length of the space of interaction (6 cm) is insufficient for an SR pulse to be generated. At long electron bunches, the peak power saturates, which is consistent with the results of our theoretical analysis. In the latter case, extra oscillation spikes (Fig. 11) that signal the transi-

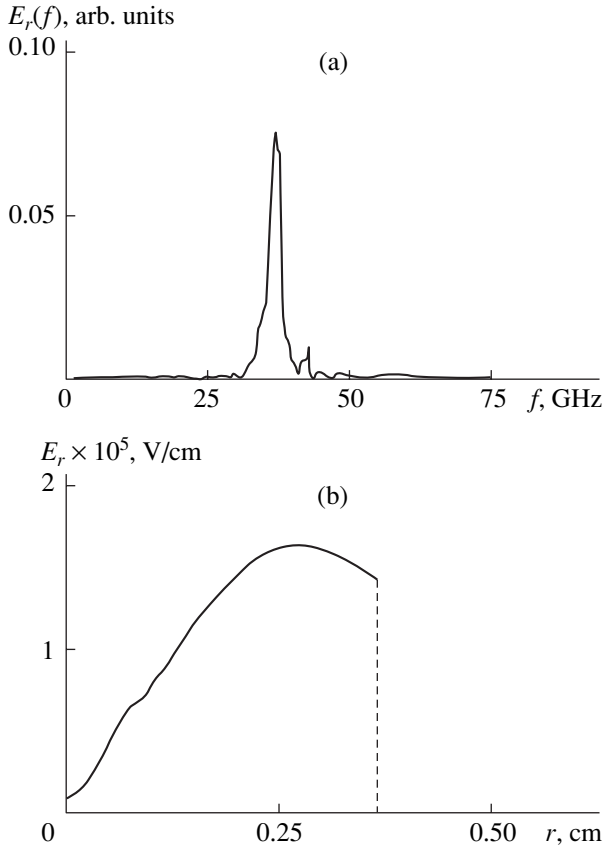


Fig. 9. (a) SRP spectrum and (b) transverse distribution of the radial component of the electric field (simulation). U_{max} = 180 kV.

tion to stochastic self-modulation in backward-wave tubes [1] appear.

Using the KARAT code, we also simulated the SR pulse generation when the amplitude of an accelerating voltage pulse at the cathode was increased to 290 kV with its shape retained. In this case, the energy of the accelerated electrons at the exit from the cathodic space increased to 300 keV with a simultaneous increase in the amplitude of the self-consistent current to 2 kA. Accordingly, the peak power of the SR pulses grew to 150 MW (Fig. 8b) in accordance with the experimental data presented below.

It should be stressed that the SR pulse peak power can be increased further by optimizing the accelerating voltage pulse shape. An optimized accelerating voltage pulse is depicted in Fig. 12a. Here, as before, the electron energy is 300 keV and the self-consistent current is about 2 kA but the SR pulse peak power rises to 400 MW (Fig. 12b). The instantaneous electronic efficiency is as high as 80% in this case.

EXPERIMENTAL DATA

When observing superradiance in a periodic slow-wave system, we considered two experimental conditions. In the first series, the accelerating voltage amplitude was 180 kV, which corresponds to an injection current of 1.2 kA. In the second one, the accelerating voltage was increased to 290 kV, which leads to an increase in the current amplitude to 2 kA.

A typical SR pulse waveform for the first case is shown in Fig. 13a. The half-height width of the pulse is about 300 ps, while the leading edge of the pulse is as short as 200 ps. It is important that the shape of the SR pulse and its peak power did not change under optimal conditions as the length of the slow-wave system varied from 6 to 10 cm. This coincides with the results of the KARAT-based numerical simulation, which also indicates that the 6-cm length suffices to generate the SR pulse under the optimum conditions.

The spectral measurements performed with a set of waveguide filters having different cutoff frequencies showed that a center radiation frequency of about 39 GHz corresponds to the excitation of the TM_{01} mode. The directivity diagram of the radiation had a

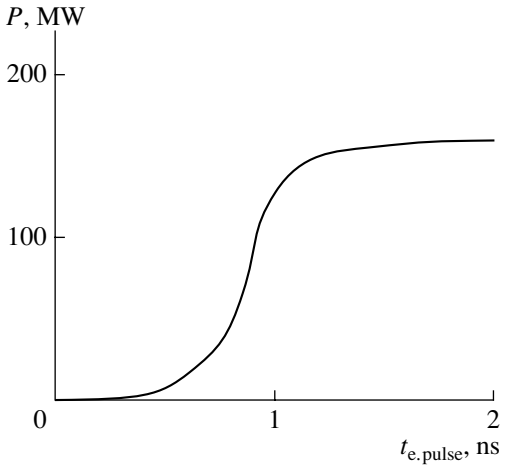


Fig. 10. SR pulse peak power vs. electron bunch duration (simulation). U_{max} = 180 kV.

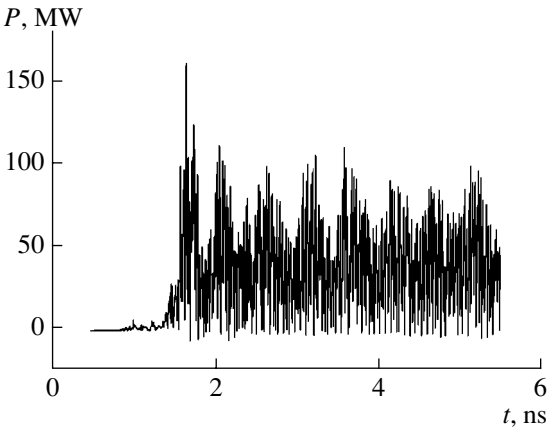


Fig. 11. Radiation in the form of a random train of pulses for a long ($t_{e.pulse}$ = 8 ns) electron bunch (simulation).

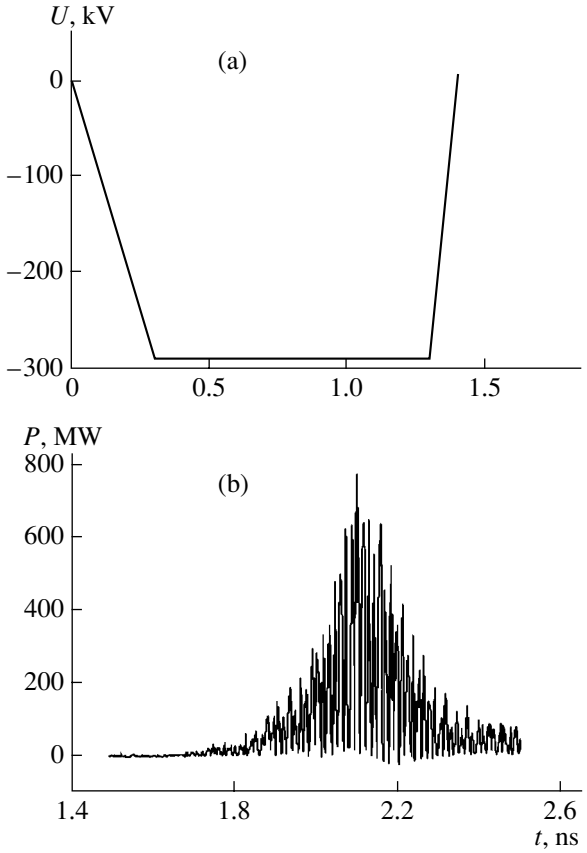


Fig. 12. (a) Optimized shape of the accelerating voltage pulse and (b) associated SR pulse with a peak power of 400 MW (simulation).

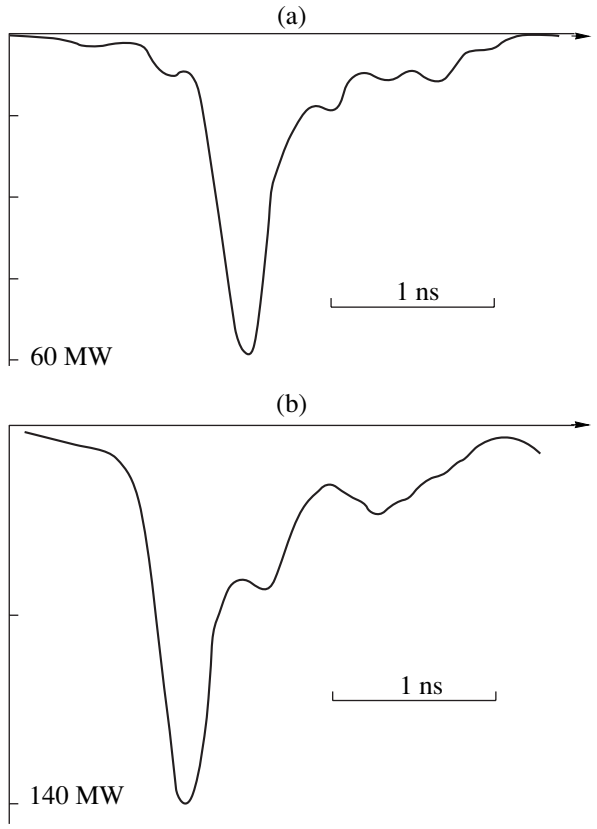


Fig. 13. SR pulse shapes for different accelerating voltage amplitudes (experiment). U_{max} = (a) 180 and (b) 290 kV.

central minimum typical of this mode. Integrating the detector readings over the directivity diagram yields the peak power at a level of 60 MW in this case.

The high level of the peak power was confirmed by the glow of a neon-tube indicator, as well as by air breakdown when the radiation was focused by a parabolic mirror or when the wave was concentrated in a conical tapered waveguide.

As was noted above, the SR effect features the quadratic dependence of the peak power on the number of particles in the bunch. As the duration of the electron pulse increases, this dependence must hold for as long as restriction (1) is met. For the Pierce parameter C = 0.10–0.15, group velocity 0.3c, translational velocity 0.7c, and radiation frequency 39 GHz, the quadratic dependence must be observed for current pulse widths up to 1 ns. The experimental dependence of the peak power on the current pulse width was constructed by varying the accelerating voltage pulse duration with the voltage amplitude and, hence, the peak current remaining unchanged. As follows from Fig. 14a, the power does grow roughly as the current pulse duration squared in the interval 0.5–1 ns. For current pulse durations below 0.5 ns, the length of the space of interaction is apparently insufficient for the SR pulse to be generated. For current pulses longer than 1.2 ns, the amplitude of the first microwave pulse stops growing and the second and then subsequent pulses are generated.

For this interval of the current pulse durations, the dependence of the peak power on the total bunch charge squared is shown in Fig. 14b. The total bunch charge is found by integrating the bunch current over time. It is seen that the curve can be approximated by a linear function with fairly good accuracy. Hence, most electrons radiate coherently. Yet, it should be noted that the peak power tends to zero faster than the total charge of the bunch. This can be explained by the fact that both the leading and the trailing edges of the bunch contain slow electrons, which cannot synchronously interact with the wave. As the electron pulse becomes shorter, the relative number of such electrons grows and the radiation power rapidly drops although the total charge of the bunch is still appreciable.

At the next stage of the experiments, the subnanosecond high-voltage generator of the accelerator was provided with an additional energy-compression system [18]. With this configuration, the accelerating voltage at the cathode increased to 290 kV and the current of the bunch was 1.5–2.0 kA. For 300-ps-wide pulses, the SR pulse peak power grew to 140–150 MW

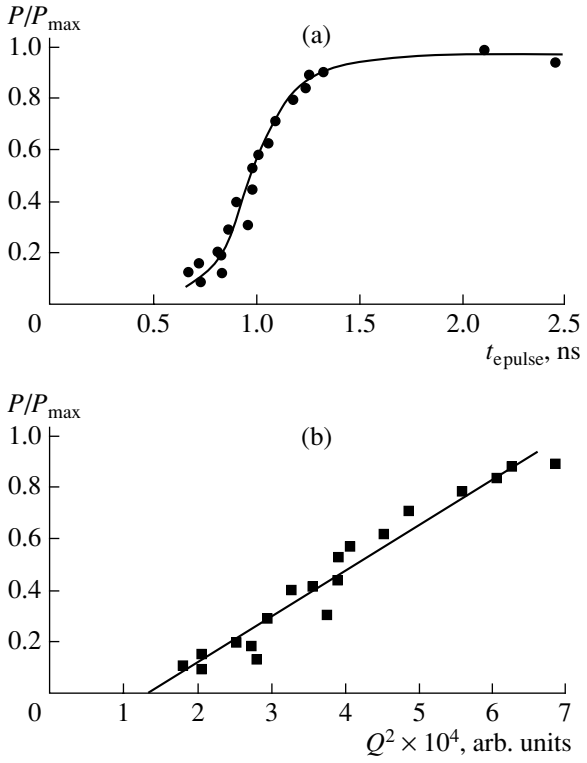


Fig. 14. SR pulse peak power vs. (a) electron bunch duration and (b) total bunch charge squared (experiment).

(Fig. 13b), which is consistent with the KARAT-based simulation (Fig. 8b).

CONCLUSION

Thus, we performed the theoretical investigation of SR pulse generation when electron bunches rectilinearly move in a periodic slow-wave system and succeeded in generating electromagnetic pulses as short as ≈300 ps with a power level as high as 140 MW. It was shown that the stable and reproducible generation of such pulses with a repetition rate of 25 Hz is a possibility. Therefore, it can be said of a new class of pulsed sources of electromagnetic radiation. Their natural applications can be plasma diagnostics with a high temporal resolution and location.

In this work, SR pulse generation was studied at frequencies near 40 GHz. To date, however, early experiments on the observation of the SR effect at frequencies near 75 and 150 GHz have been carried out [19]. The further extension of these studies into the centimeter range seems to be promising. This can be done by appropriately scaling the slow-wave system parameters and electron pulse durations.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 98-02-17308 and 01-02-17029) and the Research Program of the Russian Federation “Physics of Microwaves” (project no. 1.13)

REFERENCES

1. N. S. Ginzburg, S. P. Kuznetsov, and T. M. Fedoseeva, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **21**, 1037 (1978).
2. N. S. Ginzburg and S. P. Kuznetsov, in *Relativistic High-Frequency Electronics. Problems of Increase of Power and Radiation Frequency* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1981), pp. 101–144.
3. V. N. Shevchik and D. I. Trubetskov, *Analytic Methods of Computation in Microwave Electronics* (Sov. Radio, Moscow, 1970).
4. N. S. Ginzburg, Yu. V. Novozhilova, and A. S. Sergeev, Pis’ma Zh. Tekh. Fiz. **22** (9–10), 39 (1996) [Tech. Phys. Lett. **22**, 359 (1996)].
5. R. Bonifacio, C. Maroli, and N. Piovella, Opt. Commun. **68**, 369 (1988).
6. R. Bonifacio, N. Piovella, and B. W. J. McNeil, Phys. Rev. A **44**, R3441 (1991).
7. N. S. Ginzburg, Pis’ma Zh. Tekh. Fiz. **14**, 440 (1988) [Sov. Tech. Phys. Lett. **14**, 197 (1988)].
8. N. S. Ginzburg and A. S. Sergeev, Opt. Commun. **91**, 140 (1992).
9. N. S. Ginzburg, A. S. Sergeev, Yu. V. Novozhilova, *et al.*, Nucl. Instrum. Methods Phys. Res. A **393**, 352 (1997).
10. N. S. Ginzburg, Yu. V. Novozhilova, A. S. Sergeev, *et al.*, Pis’ma Zh. Tekh. Fiz. **24** (18), 7 (1998) [Tech. Phys. Lett. **24**, 709 (1998)].
11. É. B. Abubakirov, N. S. Ginzburg, N. F. Kovalev, and M. I. Fuks, Radiotekh. Élektron. (Moscow) **34**, 1058 (1989).
12. A. S. Elechaninov, S. D. Korovin, G. A. Mesyats, *et al.*, in *Proceedings of the 6th International Conference on High-Power Particle Beams (BEAMS-86), Kobe, 1986*, p. 552.
13. V. G. Shpak, M. I. Yalandin, N. S. Ginzburg, *et al.*, Dokl. Akad. Nauk **365**, 50 (1999) [Dokl. Phys. **44**, 143 (1999)].
14. N. S. Ginzburg, A. S. Sergeev, Yu. V. Novozhilova, *et al.*, Phys. Rev. E **60**, 3297 (1999).
15. N. F. Kovalev, Élektron. Tekh., Ser. 1: Élektron. SVCh, No. 3, 102 (1978).
16. G. A. Mesyats, V. G. Shpak, S. A. Shunailov, *et al.*, in *Proceedings of the 9th International Pulsed Power Conference, Albuquerque, 1993*, p. 835.
17. G. A. Mesyats, V. G. Shpak, S. A. Shunailov, and M. I. Yalandin, Proc. SPIE **2154**, 262 (1994).
18. V. G. Shpak, M. R. Ulmascoulov, S. A. Shunailov, and M. I. Yalandin, in *Digest of Technical Papers of the 12th IEEE International Pulsed Power Conference, Monterey, 1999*, Vol. 2, p. 692.
19. M. I. Yalandin, V. G. Shpak, S. A. Shunailov, *et al.*, IEEE Trans. Plasma Sci. **28**, 1615 (2000).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 1, 2002, pp. 88–91. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 92–95.

ELECTRON AND ION BEAMS, ACCELERATORS

A Two-Regime Probe-Forming System for Modern Nuclear Nanoprobes

S. A. Lebed

Institute of Applied Physics, National Academy of Sciences of Ukraine, ul. Petropavlovskaya 58, Sumy, 44030 Ukraine

e-mail: salmp1@yahoo.com

Received May 10, 2001

Abstract—This paper describes the long version of an optimized probe-forming system based on the divided Russian quadruplet of magnetic quadrupole lenses. The calculations include dominant intrinsic and parasitic lens aberrations. The feature of the system is a two-regime operation. This new versatile system is promising for modern nuclear nanoprobes. The system resolutions expected are presented. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Presently, about 60 scanning nuclear microprobes are in operation world-wide. More than twenty of them use probe-forming systems (PFSs) based on the Russian quadruplet of magnetic quadrupole lenses [1–8]. In the image plane, this system behaves like an axisymmetric lens [1]. This means that the Russian quadruplet produces the point image of a point object and provides equal demagnifications in two mutually perpendicular planes. As has been shown [2, 3], the divided Russian quadruplet offers this property in two regimes of lens excitation. This feature was experimentally observed by the author on the Cracow nuclear microprobe with the short (the total length $l = 2.3$ m) PFS [4]. In the first regime of lens excitation the system has a positive demagnification D. This regime with $D \approx 17$ is used in the Cracow microprobe [2–5]. The second regime (with $|D| < 2$) is impractical in the case of a short PFS. However, calculations show that it can be used as the second working regime if the probe has a long ($l \geq 8$ m) PFS.

The objective of this paper is to demonstrate that the optimized long two-regime probe-forming system based on the divided Russian quadruplet is promising for modern nuclear nanoprobes.

1. FORMULATION OF THE PROBLEM

The space position and the path of a particle in the PFS can be specified by the coordinates (x, y, z, θ, ϕ), where θ and ϕ are the instantaneous particle path slope angles projected onto the xz and yz planes, respectively. The coordinates of the particle in the image plane can be written in the form of polynomial functions [6]:

$$x_i = A_0 + A_1 x_0 + A_2 y_0 + A_3 \theta_0 + \ldots + A_n v_1^j v_2^k v_3^m + \ldots, \quad (1)$$

where the subscript 0 denotes the object coordinates and v_n is the system variable.

The order of A_n is then $j + k + m$. The aberration coefficients A_n can be expressed through the basic intrinsic and parasitic lens aberration coefficients (see Table 1). For a perfect PFS, the image is the demagnified version of an object: $x_i = x_0/D_x$, $y_i = y_0/D_y$ for any θ_0 and ϕ_0. Additional terms in expression (1) degrade the image; therefore, their effect must be minimized.

We will discuss the second approach to PFS optimization [3]. The idea is to minimize the proton beam spot size on the target for a given emittance E of the entering beam given by $E = 4(d_a\Theta)^2$, where d_a is the diameter of the object aperture and Θ is the half-angle of beam divergence after passing the angular collimator (Fig. 1). The target is assumed to be disposed in the image plane.

The parameters given are (see Table 1 and Fig. 1) lens dimensions, lens spacing s in the doublet, working distance g, ion source brightness b, microbeam emittance E, proton energy W, and spread δ of the proton momenta in the beam.

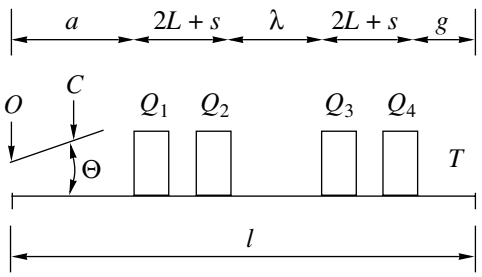


Fig. 1. Geometry of the probe-forming system based on the divided Russian quadruplet: O, object diaphragm; C, angular collimator; T, target; and Q_1–Q_4, quadrupole lenses.

1063-7842/02/4701-0088$22.00

Table 1. Physical parameters of the optimized two-regime PFS

Regimes	First	Second
Total length l, cm	950.0	950.0
Effective quadrupole length L, cm	6.4	6.4
Bore radius, cm	0.635	0.635
Object distance a, cm	867.6	867.6
Spacing between the doublets λ, cm	35.2	35.2
Working distance g, cm	15.0	15.0
Quadrupole field B_1, T	0.338	−0.087
Quadrupole field B_2, T	−0.241	0.113
Quadrupole field B_3, T	0.241	−0.113
Quadrupole field B_4, T	−0.338	0.087
Demagnification (dimensionless)		
D_x	99	–15
D_y	99	–15
Rotational aberration, μm/mrad^2		
$\langle x/\phi\rho_1\rangle$	22	–39
$\langle x/\phi\rho_2\rangle$	10	46
$\langle x/\phi\rho_3\rangle$	64	–11
$\langle x/\phi\rho_4\rangle$	−96	4
Chromatic aberration, μm/mrad/%		
$\langle x/\theta\delta\rangle$	–2172	120
$\langle y/\phi\delta\rangle$	−623	212
Spherical aberration, μm/mrad^3		
$\langle x/\theta^3\rangle$	50 750	−43
$\langle x/\theta\phi^2\rangle$	5100	–220
$\langle y/\theta^2\phi\rangle$	5100	–220
$\langle y/\phi^3\rangle$	2310	–203
Parasitic aberration sensitivity	$\langle x/\theta^2 S3\rangle$	$\langle y/\phi\theta S2\rangle$
Largest sextupole term, μm/mrad^2/%	–30 590	−878
Largest octupole term, μm/mrad^3/%	$\langle x/\theta^3 O3\rangle$	$\langle y/\phi^3 O2\rangle$
	–243 960	–2110

Note: $S1$, $O1$, etc. stand for the percentage of the parasitic sextupole and octupole contaminations from the pole tip field, respectively, in quadrupole lens 1, etc [7, 11].

The conventional VdG HVEC (δ = 0.05%) [2, 4] and novel ultrastable Singletron (δ = 0.001%) [7] accelerators allow the probe operation with a paraxial (Θ ≤ 0.2 mrad) high-brightness ($b \approx 10$ and 18 A m^{-2} rad^{-2} eV^{-1}, respectively) proton beam.

The variable parameters are as follows: PFS total length ($5 \le l \le 10$ m), spacing λ between the doublets, object diaphragm diameter ($1 \le d_a \le \sim 25$ μm), and quadrupole lens excitations ($0.1 \le k_i \le 1.0$, where $i = 1,\ldots, 4$). The value of the ion beam current I at the target can be determined from the formula $I = EbW$. The beam spot size on the target, $d = \max\{d_x, d_y\}$, is calculated as a function of the length l and spacing λ. Unlike [3], in this paper, the calculation was performed simultaneously at two stigmatic points of proton beam focusing and for two current modes. This made it possible to find a two-regime PFS in which one regime was optimized for the operation in the low-current ($I \le 1$ pA) mode with high spatial resolution, while the second regime, for the operation in the high-current ($I \ge 100$ pA) mode with a submicron resolution. Examples of the low-current modes are the modes of scanning transmission ion microscopy and the ion-beam-induced charge mode. High-current modes are used in characteristic X-ray radiation technique and Rutherford forward- and backscattering spectrometry.

Table 2. Calculated resolutions of nanoprobes based on the optimized two-regime PFS

Mode	Low-current		High-current	
accelerator type	VdG HVEC	singletron	VdG HVEC	singletron
Object aperture diameter d_a, μm	1.0	1.0	10.0	10.0
Proton beam divergence Θ, mrad	0.0006	0.0006	0.05	0.037
Beam size on the target $d_x \times d_y$ in the first regime, μm	0.03×0.02	0.012×0.011	$\sim 25 \times 4$	$\sim 14 \times 2$
Beam size on the target, $d_x \times d_y$ in the second regime, μm	0.07×0.07	0.07×0.07	0.85×0.95	0.70×0.73

2. RESULTS

The numerical simulation of the PFS was performed using the SFS [8], PRAM [9], and TRANSPORT [10] codes. The calculations included all dominant intrinsic (chromatic and spherical) and parasitic (rotational, sextupole, and octupole) lens aberrations.

Present-day microprobe quadrupole single lenses have relatively low values of the sextupole and octupole parasitic contaminations from the pole tip field (0.05–0.3%) [11, 12].

Table 1 lists the physical parameters of the optimized two-regime PFS. The system is to be used as a new long (9.5 m) PFS version in the Cracow nuclear nanoprobe. Therefore, the geometry of the quadrupole doublets of the magnetic quadrupole lenses from the Micro Analytical Research Center (Melbourne University, Australia) [4] were employed as the input in the calculations. The neighboring (in the z direction) pole tip pairs of the doublets (Q_1–Q_2 and Q_3–Q_4, Fig. 1) and their common yokes were cut out of one piece of commercially pure iron with tolerances of ≈5 μm by means of an electro-erosion machine. All these parts were bolted together into a doublet with tolerances of 10–15 μm in the x, y, and z directions. These design features help to eliminate translational, tilt, and rotational misalignments [6] in the quadrupole lenses and in the doublet. Moreover, both PFS doublets were mounted on special 3D-adjustable mechanical devices to precisely align the doublets by the beam spot on the quartz target [4]. In this way, the spot broadening due to the first- and second-order parasitic lens aberrations was considerably decreased.

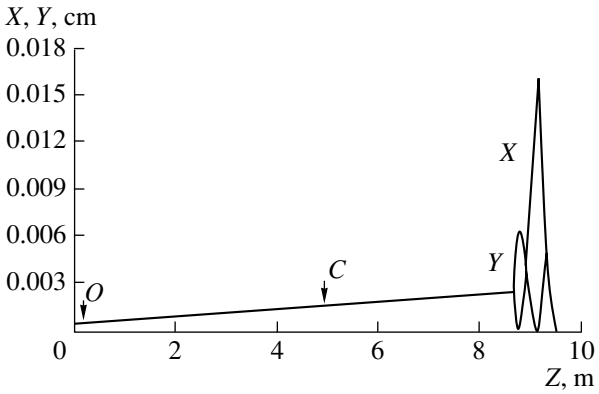


Fig. 2. Beam envelopes along the proton path that are calculated for the low-current mode of the first regime.

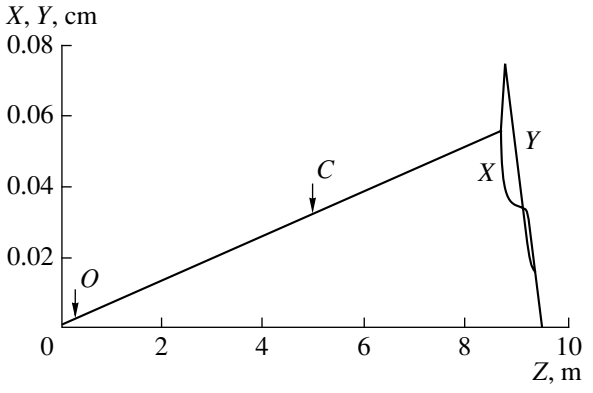


Fig. 3. Beam envelopes along the proton path that are calculated for the high-current mode of the second regime.

The novel post-lens ion beam scanning system [4, 13] is an important advantage of the Cracow microprobe. This short (in the z direction) ferrite-core electromagnetic system enables high-precision scanning of MeV beams within a short working distance (15 cm).

Table 2 presents the expected beam spot sizes on the target that were calculated for the optimized two-regime PFS in both current modes for the conventional and novel (ultrastable) accelerators. The beam envelopes along the proton path for these regimes and modes are shown in Figs. 2 and 3. The calculations were performed for the proton energy W = 2.5 MeV with 0.3% sextupole and octupole parasitic components included in all the lenses.

The author also performed the numerical optimization of alternative PFSs based on a magnetic quadrupole doublet or a triplet under similar laboratory conditions. It was found that the system based on the optimized divided Russian quadruplet provides the highest currently available demagnification of a modern nuclear nanoprobe in both directions ($|D| = |D_x| = |D_y| >$ 90). This advantage of such a PFS over alternative systems is the most pronounced when the probe operates in the low-current mode.

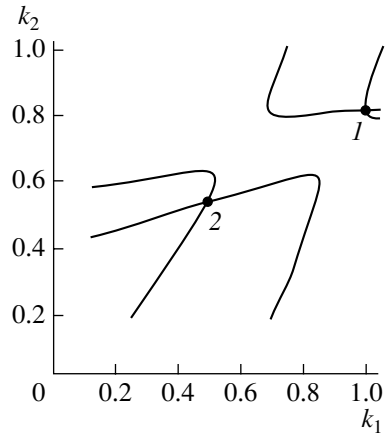


Fig. 4. Excitation k_2 of the second and third lenses as a function of the excitation k_1 of the first and fourth lenses in the divided Russian quadruplet.

3. DISCUSSION

As can be seen (Table 1), the optimized two-regime PFS has the relatively short spacing between the doublets in the divided Russian quadruplet (λ = 35.2 cm). That is why the first and second regimes exhibit the ion–optical properties similar to those observed under the working conditions of the divided and classical Russian quadruplets, respectively.

The first regime, characterized by intermediate beam crossovers (Fig. 3) and a positive symmetric demagnification ($D = D_x = D_y = 99$), is achieved at higher lens excitations ($k_1 = 0.98$, $k_2 = O.82$; point *1* in Fig. 4) in the quadruplet than the second regime ($k_1 = 0.5$, $k_2 = 0.56$; point *2* in Fig. 4). Because of the large D, the first regime allows the nanoprobe to have ultimately high spatial resolutions (down to 12 nm for the ultrastable accelerator, Table 2). However, the first regime is strongly sensitive to intrinsic and parasitic lens aberrations (Table 1). This significantly broadens the beam spot size in the high-current mode (Table 2).

The second PFS regime does not have intermediate beam crossovers (Fig. 3). Its basic advantage is low sensitivity to the above aberrations (Table 1). This regime with compromise ($D = D_x = D_y = -15$) provides a rather high resolution (70 nm) in the low-current mode and a submicron resolution in the high-current mode even in the case of conventional accelerators (Table 2).

It is seen in Fig. 4 that, for a transition from one PFS regime to another, it is necessary and sufficient to accurately choose a pair of lens excitations. In practice, the quality (stigmatism) of such a transition can easily be controlled by visually monitoring the shape of a microbeam on the quartz target.

CONCLUSION

A new optimized two-regime PFS for modern nanoprobes that is based on the divided Russian quadruplet of magnetic quadrupole lenses is suggested. Its basic advantages are (i) extremely high spatial resolution (down to 12 nm) in the low-current modes in the presence of intermediate beam crossovers, (ii) submicron resolution in the high-current modes in the absence of intermediate beam crossovers (even for a conventional accelerator), and (iii) easy switching from the first to the second working regime (and in reverse) without the expensive readjustment of the beam line equipment.

REFERENCES

1. A. D. Dymnikov and S. Ya. Yavor, Zh. Tekh. Fiz. **33**, 851 (1963) [Sov. Phys. Tech. Phys. **8**, 639 (1964)].
2. S. Lebed, Nucl. Instrum. Methods Phys. Res. B **155**, 322 (1999).
3. S. Lebed, Nucl. Instrum. Methods Phys. Res. B **152**, 145 (1999).
4. S. Lebed, M. Cholewa, Z. Chioch, *et al.*, Nucl. Instrum. Methods Phys. Res. B **158**, 44 (1999).
5. V. Brazhnik, S. Lebed, W. Kwiatek, *et al.*, Nucl. Instrum. Methods Phys. Res. B **130**, 104 (1997).
6. G. W. Grime and F. Watt, *Beam Optics of Quadrupole Probe-Forming Systems* (A. Hilger, Bristol, 1984), p. 273.
7. D. J. W. Mous, R. G. Haitsma, T. Butz, *et al.*, Nucl. Instrum. Methods Phys. Res. B **130**, 31 (1997).
8. V. Brazhnik, S. Lebed, V. I. Miroshnichenko, *et al.*, Nucl. Instrum. Methods Phys. Res. B **171**, 558 (2000).
9. M. Breese, D. Jamieson, and P. King, *Materials Analysis with a Nuclear Microprobe* (Wiley, New York, 1996).
10. K. Brown, SLAC Rep. SLAC-91 (1977).
11. D. Jamieson and G. Legge, Nucl. Instrum. Methods Phys. Res. B **30**, 235 (1988).
12. D. Jamieson, J. Zhu, Y. Mao, *et al.*, Nucl. Instrum. Methods Phys. Res. B **104**, 86 (1995).
13. V. Khomekno, S. Lebed, S. Mordik, *et al.*, Nucl. Instrum. Methods Phys. Res. B **130**, 86 (1997).

Translated by the author

Technical Physics, Vol. 47, No. 1, 2002, pp. 92–96. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 96–99.

SURFACES, ELECTRON AND ION EMISSION

On the Relationship between the Ionization Potential and the Work Function: Metals

S. Yu. Davydov

St. Petersburg State Electrotechnical University, St. Petersburg, 197376 Russia

Received January 26, 2001

Abstract—A relationship is analyzed between the ionization energy of atoms I and the work function ϕ of a metal composed of these atoms. The transition energy $\Sigma = I - \phi$ is assumed to be the sum of kinetic K and coulombic C contributions. Contribution K is calculated in the framework of a model of uniform gas of quasi-free electrons, and C is determined from experimental values of Σ. Calculations performed for a wide range of metals have shown that the dimensionless coefficients determining the Coulombic contribution C differ insignificantly between various groups of metals. The relationships obtained have been used for determining the work function of binary alloys. © *2002 MAIK "Nauka/Interperiodica".*

Attempts to calculate the work function ϕ of pure metals having a relatively simple and well-understood electronic structure and nearly spherical Fermi surface (which include, for example, alkali metals) had been undertaken back in the 1930s by Frenkel [1] and also by Tamm and Blochinzev [2]. Later, a more rigorous quantum mechanical theory was developed by Wigner and Bardeen [3, 4]. They were able to present the exchange and correlation energy of a uniform electronic gas as a function of electron density n_e, thus making possible calculation of the work function ϕ by way of differentiating the energy with respect to n_e. The concepts introduced in these works had a great impact on all subsequent theoretical studies. An important stage in the development of the work function theory was the use of a formalism of density functional [5]. It has been shown that the work function is a sum of two components, those of bulk and surface, with the exchange-correlation interaction contributing both to the chemical potential of the metal bulk and to the surface barrier. In particular, the effect of image forces can be treated as a removal of some part of the correlation energy when an electron leaves the metal surface.

This formalism, however, gives adequate description of only limited classes of metals, such as alkali metals and other *s–p*-metals, whereas the greatest technological interest is drawn to refractory substrates of *d*-metals. The *d*-metals (as *f*-metals) have in their electron spectra resonance *d*(*f*)-bands causing a sharp change of the density of energy states in the substrate over a relatively narrow range of energies, which is difficult to describe in the framework of the density functional formalism. Therefore, the results obtained with this method, while yielding useful qualitative information, can only reveal general tendencies of the dependence of ϕ on the electron gas density.

Complexity of the quantum-mechanical calculations of the work function stimulated the development of numerous empirical and semiempirical estimates, attempting to relate the magnitude of ϕ with some characteristics of metals and their constituent atoms (see, e.g., [6]). In this work, an attempt is made at finding, with the use of a semiempirical method, a correlation between the ionization potential of the atoms comprising a given metal and its work function.

FROM ATOM TO METAL

Energywise, the basic characteristic of an atom is its ionization potential I, i.e., the energy required for removing an outer electron of the atom to infinity. When atoms join together and form a crystal, the atomic levels shift and expand into bands. For a metal, the characteristic analogous to the ionization potential is the work function ϕ, or the energy needed to remove an atom from the Fermi level to infinity. Somewhat loosely, it can be stated that when atoms assemble into a crystal, the outer electron goes from level I to the Fermi level ϕ. The energy necessary for the transition is

$$\Sigma = I - \phi. \quad (1)$$

Because all interactions of coulombic nature in a crystal depend on the distance r between interacting particles (electrons, nuclei) as r^{-1}, a natural unit of the Coulomb interaction in a crystal should be a quantity e^2/d, where e is the electron charge and d is the distance between two nearest-neighbor atoms in the crystal. As atoms join to form a crystal, apart from the Coulomb forces, the so-called volumetric force comes into play, originating because of the system's gain in kinetic energy accompanying delocalization (collectivization) of electrons. The last statement is easily understood proceeding from the uncertainty relation. Indeed, in an atom, the uncertainty of the electron's position is deter-

1063-7842/02/4701-0092$22.00

Table 1. Calculation results of coulombic C and kinetic K contributions to the transition energy Σ and of the coefficient D for alkali- ($Z_s = 1$) and alkali-earth ($Z_s = 2$) metals, metals of the zinc group ($Z_s = 2$), and metals of the aluminum group ($Z_p = 1$)

Metal	Σ, eV	K, eV	C, eV	D	K/C
Li	3.01	–1.42	4.43	0.93	–0.32
Na	2.79	–0.97	3.76	0.96	–0.26
K	2.12	–0.63	2.75	0.87	–0.21
Rb	2.02	–0.55	2.57	0.87	–0.18
Cs	2.08	–0.47	2.55	0.93	–0.18
Be	5.40	3.35	2.05	0.32	1.63
Mg	4.00	1.61	2.39	0.53	0.67
Ca	3.31	1.06	2.25	0.62	0.47
Sr	3.34	0.89	2.45	0.73	0.36
Ba	2.72	0.66	2.06	0.62	0.32
Zn	5.15	2.34	2.81	0.52	0.83
Cd	4.89	1.86	3.03	0.63	0.61
Hg (*hcp*)	5.92	1.82	4.10	0.86	0.44
Al	1.74	–1.28	3.02	0.60	–0.42
Ga (*fcc*)	2.04	–1.76	3.80	0.64	–0.46
(*bcc*)	2.00	–2.18	4.22	0.72	–0.52
In (*fcc*)	1.99	–0.99	2.98	0.67	–0.33
(*bcc*)	1.99	–1.23	3.22	0.73	–0.38
Tl	2.41	–0.87	3.28	0.79	–0.27

mined by the radius R of the electron's orbit, and the uncertainty of its momentum is $\Delta p \sim \hbar/R$ ($\hbar$ being the Plank constant). In a metal, the uncertainty of a coordinate tends to infinity because an electron in the conduction band can travel anywhere within its bounds. This lowers the kinetic energy of an electron in a metal, and this effect is the basis of the metallic bond. Because in a quasi-free electron gas the Fermi energy is $\varepsilon_f = \hbar^2 k_f^2/2m$ (m being the electron mass) and the Fermi vector is $k_f \sim d^{-1}$, it is easy to see that an appropriate unit of measurement for the contribution of volumetric forces to energy Σ is the quantity $\hbar^2/md^2$ [7]. Then the transition energy Σ can be represented in the form

$$\Sigma = C + K, \quad C = D(e^2/d), \quad K = B(\hbar^2/md^2), \quad (2)$$

where C and K are the coulombic and kinetic contributions, respectively, and D, B are dimensionless coefficients.

For any multielectron system, the most challenging problem is that of calculating the exchange-correlation interaction, the contribution of exchange forces being determined mainly by the Fermi–Dirac statistics. Following Wills and Harrison [8, 9], we assume that the Fermi energy of a metal is determined by quasi-free electrons, the energy of which can be expressed as

$$E(k) = \varepsilon_s - \frac{3}{4}\frac{\hbar^2\pi^2}{md^2} + \frac{\hbar^2 k^2}{2m}, \quad (3)$$

where k is a quasi wave vector and ε_s is the energy of an external s-level forming the conduction band and measured from the level of a vacuum.

Now, in the framework of the one-electron approximation, we can juxtapose to the magnitude of the ionization potential the quantity $(-\varepsilon_s)$ and with the work function ϕ the quantity $E_f \equiv E(k_f)$ taken with the opposite sign, i.e., $\Sigma = (E_f - \varepsilon_s)$. On the other hand, $\Sigma = K$ because in the one-electron approximation $C = 0$. From expression (3), we find

$$B = \frac{1}{2}(\xi^2 - 3\pi^2/2), \quad \xi = aZ_s^{1/3}, \quad (4)$$

where Z_s is the number of electrons in the s-band, $a = (3\sqrt{2}\,\pi^2)^{1/3}$ for *fcc* and *hcp* lattices, and $a = (9\sqrt{3}\,\pi^2/4)^{1/3}$ for *bcc* lattices [8, 9].

As follows from expression (4), the magnitude of B depends only on the lattice type and the band occupation number Z_s, which varies between 1 and 2 and is proportional to $Z_s^{2/3}$. At $Z_s = 1$, $B = -1.37$ for *fcc* and *hcp* structures and $B = -1.71$ for *bcc* structures; at $Z_s = 2$ we obtain, correspondingly, $B = 2.17$ and 1.64, with B becoming zero at $Z_s^* = 1.36$ for *fcc* and *hcp* lattices and at $Z_s^* = 1.48$ for *bcc* lattices. Coefficient B changing sign from minus to plus corresponds to the change of

Table 2. Same as Table 1 for transition metals (Z_s = 1.5)

Metal	Σ, eV	*K*, eV	*C*, eV	*D*	*K*/*C*
Ss	3.26	0.36	2.90	0.66	0.12
Ti	2.87	0.45	2.42	0.49	0.19
V	2.62	0.07	2.55	0.46	0.03
Cr	2.19	0.08	2.11	0.37	0.04
Mn (*fcc*)	3.60	0.34	3.26	0.51	0.10
(*bcc*)	3.60	0.10	3.50	0.55	0.03
Fe	3.59	0.08	3.51	0.60	0.02
Co	3.45	0.60	2.85	0.49	0.21
Ni	3.14	0.61	2.53	0.44	0.24
Cu	3.33	0.58	2.75	0.49	0.21
Y	2.92	0.30	2.62	0.65	0.11
Zr	2.73	0.38	2.35	0.52	0.16
Nb	2.77	0.06	2.71	0.54	0.02
Mo	2.79	0.06	2.73	0.51	0.02
Tc	–	0.51	–	–	–
Ru	2.77	0.54	2.23	0.41	0.24
Rh	2.71	0.52	2.19	0.41	0.24
Pd	3.54	0.50	3.04	0.58	0.16
Ag	3.27	0.45	2.82	0.57	0.16
Lu	2.10	0.32	1.78	0.42	0.18
Hf	3.27	0.39	2.88	0.63	0.14
Ta	3.77	0.06	3.71	0.74	0.02
W	3.44	0.06	3.38	0.64	0.02
Re	2.88	0.50	2.38	0.45	0.21
Os	4.03	0.46	3.57	0.66	0.13
Ir	3.78	0.51	3.27	0.61	0.16
Pt	3.64	0.49	3.15	0.61	0.16
Au	4.13	0.46	3.67	0.73	0.13

sign of the difference ($E_f - \varepsilon$): in a half-filled band (Z_s = 1), the ε_s level is higher than the E_f level; in a completely filled band (Z_s = 2), the positions of the levels are reversed.

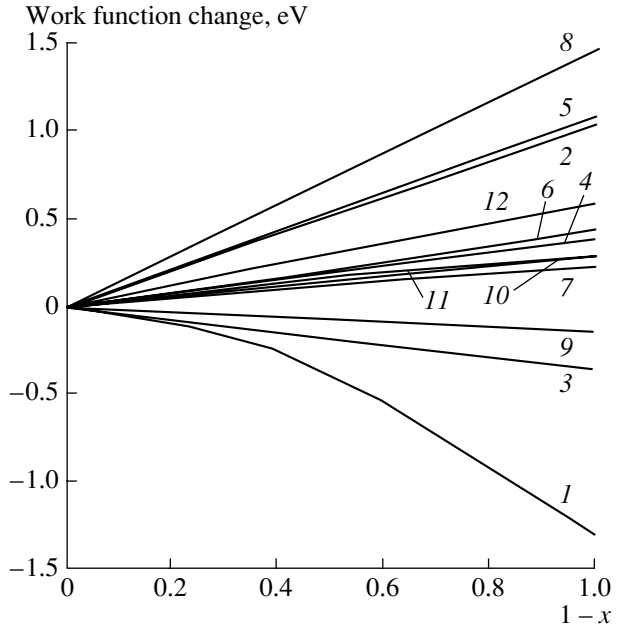
By calculating coefficient *B* and, thus, the kinetic contribution *K* of a uniform gas of noninteracting electrons to the transition energy Σ by formulas (2) and (4), we can then find the coulombic component *C* from relationship (1) using experimental data for ionization energies and work function values. Coefficient *D* calculated in this manner formally contains all corrections for both the exchange-correlation interaction and the nonuniformity of the electron gas (due to the effect of the surface)[1].

Experimental values of the transition energy Σ and the results of the calculation of the kinetic *K* and coulombic *C* contributions, their ratios, and values of the coefficient *D* are given in Tables 1–3. Data on the ionization potentials and work functions (for polycrystals) are taken from handbook [10] and the nearest-neighbor distances from [11]. Values of the *s*-band filling are chosen in the same way as in works [8, 9, 12]. So, for alkali metals Z_s = 1; for metals of Group II, Z_s = 2, and lanthanides Z_s = 1.5. In metals of the aluminum group, the outer shell only contains a single *p*-electron (here we use the same estimation as for *s*-electrons, assuming Z_s = 1).

Calculation results show that the parameters e^2/d and $\hbar^2/md^2$ define a natural energy scale because coefficients *D* and *B* take values on the order of unity. For metals of the same group, the coulombic coefficients *D* do not differ very much, especially in the transition metal series. For alkali metals an arithmetical average of the coefficients $\bar{D} \cong 0.91$; for 3*d*-metals $\bar{D} \cong 0.51$; for 4*d*-metals $\bar{D} \cong 0.52$; for 5*d*-metals $\bar{D} \cong 0.61$; for lanthanides $\bar{D} \cong 0.62$; for the beryllium group, where the difference between the coefficients is maximal, $\bar{D} \cong 0.56$; for the zinc group, where the difference between the coefficients is also great, $\bar{D} \cong 0.67$; and for the aluminum group $\bar{D} \cong 0.69$. It is seen that the scatter between the groups is not very significant.

It is noteworthy that negative values of the kinetic contribution *K* have been obtained only for alkali metals and metals of the aluminum group, whose atoms have on their outermost orbital only one electron. For all metals with the exception of beryllium, the ratio $\eta \equiv K/C < 1$. In the series of alkali metals and metals of beryllium and zinc groups, this ratio is smoothly decreasing going from light to heavy elements (Table 1). As $\eta \propto d^{-1}$, this behavior is explained by the increase of the lattice constant in the same series. This statement is true for metals of the aluminum group as well: arranging these elements according to their lattice constants, we get the series Ga–Al–In–Tl. From Table 1, it follows that in the same series the ratio η is decreasing. For transition metals with *fcc* and *hcp* structure, *K*/*C* varies in a range from 0.10 to 0.24. For body-centered lattices, the ratio *K*/*C* is in a range of 0.01–0.03. Such small values of *K*/*C* for the bcc structures are

[1] Taking into account the nonuniformity of the electron gas in the form of the coulombic contribution *C* is, strictly speaking, incorrect because this factor is proportional to d^{-2}, not d^{-1}. Still, we consider that the error thus introduced in the calculations is insignificant.



Change of the work function $\Delta\phi$ of $M_{1,x} = -M_{2,(1-x)}$ alloys as a function of concentration x of component M_1: (*1*) Na–K; (*2*) Ti–Re; (*3*) Rb–Cs; (*4*) Zr–Mo; (*5*) Zr–Re; (*6*) Mo–Rh; (*7*) Mo–W; (*8*) Hf –Re; (*9*) Ta–Nb; (*10*) Cd–Bi; (*11*) Tl–Pb; (12) In–Bi. Curves *10* and *11* nearly coincide.

explained by an extremely small kinetic contribution K, because in this case the value $Z_s = 1.5$ taken in the calculation is very close to the critical value $Z_s^* = 1.48$. In the case of *d*- and *f*-metals, no unique relationship between parameter η and the lattice constant is found.

Table 3. Same as Table 1 for lanthanides ($Z_s = 1.5$)

Metal	Σ, eV	K, eV	C, eV	D	K/C
La	2.28	0.27	2.01	0.52	0.13
Ce	2.84	0.28	2.56	0.65	0.11
Pr	2.77	0.29	2.48	0.63	0.12
Nd	2.33	0.28	2.05	0.52	0.14
Pm	–	–	–	–	–
Sm (*hcp*)	2.94	0.29	2.65	0.66	0.11
(*bcc*)	2.94	0.04	2.90	0.72	0.01
Eu	3.17	0.03	3.14	0.86	0.01
Gd	3.05	0.29	2.76	0.68	0.11
Tb	2.71	0.31	2.40	0.59	0.13
Dy	2.69	0.31	2.38	0.58	0.13
Ho	2.80	0.31	2.49	0.60	0.12
Er	2.86	0.31	2.55	0.61	0.12
Tm	3.06	0.30	2.76	0.68	0.11
Yb	3.88	–	–	–	–
Lu	2.10	0.32	1.79	0.42	0.18

FROM METAL TO ALLOY

To illustrate the application of the results, consider continuous substitutional binary solid solutions $M_1 - M_2$. Suppose that the lattice constant of the solid solution is varying with concentration according to the Vegard law [13], i.e., linearly,

$$d(x) = xd_1 + (1-x)d_2, \tag{5}$$

where x and $(1 - x)$ are concentrations of the components M_1 and M_2, respectively.

Then the effective work function $\phi(x)$ of an alloy can be represented in the form

$$\begin{aligned} \phi(x) &= xI_1 + (1-x)I_2 - \Sigma(x), \\ \Sigma(x) &= [e^2/d(x)][xD_1 + (1-x)D_2]. \end{aligned} \tag{6}$$

It is easily seen that the suggested expression for $\phi(x)$ is nonlinear in the concentration. The slope of the concentrational dependence of the work function is

$$\begin{aligned} \partial\phi(x)/\partial x &= I_1 - I_2 - \partial\Sigma(x)/\partial x, \\ \partial\Sigma(x)/\partial x &= d(x)^{-2}[2xD_1d_1 - 2(1-x)D_2d_2 \\ &\quad + (1-2x)(D_1d_2 + D_2d_1)], \end{aligned} \tag{7}$$

and the second derivative is

$$\partial^2\phi(x)/\partial x^2 = [2/d(x)^2](d_1 - d_2)(d_1D_2 - d_2D_1). \tag{8}$$

From expression (8) it follows, in particular, that the second derivation does not change sign as the concentration changes; consequently, the $\phi(x)$ curve is either convex or concave. In the framework of the present model, any nonmonotonic variation of the work function in a binary system is related to some chemical transformations in the alloy. From (6) and (7), it also follows that the less the difference between parameters D and d of the alloy components, the closer the $\phi(x)$ dependence to a linear one.

Calculation results for a number of compounds are presented in the figure. Comparison of the calculation results with experimental data in [6] (see section "Alloys" and Figs. 103–113; also given are references to original works) show quite satisfactory agreement; however, it should be noted that values of the work function for components of the binary alloys given in [10] and in original works cited in [6] are sometimes at odds.

Thus, in this study, an attempt is undertaken to relate two fundamental characteristics: the ionization potential I of an atom and the work function ϕ of the metal. From calculations of the kinetic contribution K to the transition energy Σ, in the approximation of a uniform gas of quasi-free electrons, we have determined, using experimental data on Σ, the coulombic contribution C and included in it a correction for gas nonuniformity. Analysis for a large number of metals has shown that the coefficient D determining the coulombic contribution is on the order of unity and varies insignificantly

for various groups of metals. This can be considered evidence that the developed approach to describe the work function is quite adequate.

ACKNOWLEDGMENTS

This study was carried out in the framework of the federal program “Surface Atomic Structures” (project no. 4.5.99).

REFERENCES

1. J. Frenkel, Z. Phys. **49**, 31 (1928).
2. J. Tamm and D. Blochinzev, Z. Phys. **77**, 424 (1932).
3. E. Wigner and J. Bardeen, Phys. Rev. **48**, 84 (1935).
4. J. Bardeen, Phys. Rev. **49**, 653 (1936).
5. N. D. Lang, in *Solid State Physics*, Ed. by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1973), Vol. 28, pp. 225–308.
6. V. S. Fomenko, *Emission Properties of Materials: Handbook* (Naukova Dumka, Kiev, 1981, 4th ed.).
7. W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* (Freeman, San Francisco, 1980; Mir, Moscow, 1983), Vol. 2.
8. J. M. Wills and W. A. Harrison, Phys. Rev. B **28**, 4363 (1983).
9. J. M. Wills and W. A. Harrison, Phys. Rev. B **29**, 5486 (1984).
10. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meilikhov (Énergoatomizdat, Moscow, 1991; CRC Press, Boca Raton, 1997).
11. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976; Nauka, Moscow, 1978).
12. S. Y. Davydov and S. K. Tikhonov, Surf. Sci. **371**, 157 (1997).
13. B. F. Ormont, *Introduction to Physical Chemistry and Crystal Chemistry of Semiconductors* (Vysshaya Shkola, Moscow, 1968).

Translated by B. Kalinin

Technical Physics, Vol. 47, No. 1, 2002, pp. 97–101. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 1, 2002, pp. 100–104.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

The Simulation of the Interaction between Low-Intensity Laser Beams and Multilayer Scattering Biological Materials

A. Yu. Seteĭkin, M. M. Gershevich, and I. A. Ershov

Amur State University, Blagoveshchensk, 675027 Russia

e-mail: master@amursu.ru

Received April 10, 2001

Abstract—A mathematical model of interaction between low-intensity laser radiation and multilayer biological materials with allowance for scattering is put forward. Within this model, the intensity distribution during the laser irradiation is found by the Monte Carlo method and the temperature field in the biological tissue, by the method of adaptive finite elements that automatically generates an optimal element grid. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Laser radiation of various wavelength has recently found wide application in medicine. In particular, low-intensity laser radiation (LILR) is successfully used in treating many diseases [1, 2]. Unfortunately, there is no agreement in the medical community regarding the biological (therapeutic) effect of LILR. It seems most likely that laser radiation activates or inhibits different photochemical reactions in biological tissues under the conditions of synergy.

The character of laser radiation–biological tissue interaction depends on the power density of the incident light, as well as on the optical and thermophysical properties of the tissue. The radiation of He–Ne and semiconductor lasers that are in most common use in laser therapy falls into the "therapeutic" window (λ = 0.5–1.5 μm), where dynamic scattering dominates over absorption. Reliable information on the penetration depth of a laser beam, temperature field, thermal-effect area, and absorbed dose is difficult to gain because of the complexity of related experiments and the lack of rigorous analytical calculations. The known methods for computing the temperature fields with regard for scattering are inaccurate or involve complicated procedures of finding desired coefficients [3, 4].

In this work, we consider the interaction of LILR with various biological materials using the Monte Carlo method. This method adequately predicts the distribution of the radiation intensity $I(x, y, z, t)$ in the material, taking into account its multilayer structure and the finite sizes of the incident beam. From the distribution $I(x, y, z, t)$, one can find heat sources in the medium by solving the boundary-value problem of thermal conduction, derive the temperature fields $T(x, y, z, t)$ in the biological material, and determine the thermal-effect areas.

THE BASIC EQUATIONS OF THE METHOD

When simulating, a biological material is assumed to be a multilayer medium. Each of the layers is characterized by the absorption coefficient μ_a; scattering coefficient μ_s; refractive index n; and the effective absorption coefficient

$$\mu_d = \sqrt{3\mu_a(\mu_a + (1-g)\mu_s)}, \quad (1)$$

where g is the anisotropy parameter.

The values of the optical parameters are taken from the literature [3, 5] and listed in the table.

We assume that scattering is weak in all biological tissues under consideration. The value of g is easily found if the Henny–Greenstein model phase function of scattering is known [3]:

$$p(\cos\Theta) = \frac{1}{4\pi}\frac{1-g^2}{(1+g^2-2\cos\Theta)^{\frac{3}{2}}}, \quad (2)$$

$$g = \langle\cos\Theta\rangle = 2\pi\int_0^\pi p(\cos\Theta)\cos\Theta\sin\Theta d\Theta. \quad (3)$$

In the calculations, the laser beams are assumed to be collimated and strike the surface at a right angle. The beams have various initial radii and the Gaussian or

Optical parameters of the tissues (λ = 633 nm)

Tissues	μ_a, mm^{-1}	μ_s, mm^{-1}	g	n	d, μm
Epidermis	0.43	10.70	0.79	1.50	65
Upper derma	0.27	18.70	0.82	1.40	565
Capillary plexus	2.5	40	0.98	1.35	90
Lower derma	0.27	18.70	0.82	1.40	565

1063-7842/02/4701-0097$22.00

rectangular profile (such profiles are typical of beams at the exit from an optical fiber). For the Gaussian beam, the optical intensity distribution has the form

$$I(\rho, t) = I_0 \exp\left\{-2\left(\frac{\rho}{R_0}\right)^2\right\}\xi(t) \quad (4)$$

for that with a rectangular profile,

$$I(\rho, t) = \begin{cases} 0, & \rho > R_0 \\ I_0\xi(t), & \rho < R_0. \end{cases} \quad (5)$$

Here, $I_0 = 2P/\pi R_0^2$, P is the power, R_0 is the initial radius of the beam, and $\xi(t)$ is the Heaviside function: $\xi(t) = 1$ for $t > 0$ and 0 for $t < 0$.

The optical properties are assumed to be invariable for each of the layers; that is, the problem is symmetric under shift in the plane parallel to the interfaces. When using the statistical method, we approximated the laser beam by a photon flux with a large number of the particles, $N > 10^5$.

The statistical simulation of photon motion is accomplished as follows. The domain of interest (ρ, z) $(0 \le z \le t$, where $t = \sum_{i=1}^{N} t_i$ is the total thickness of the material and N is the number of the layers; $0 \le \rho \le \rho_{max}$, where $\rho_{max} > 10R_0$; a fraction of the radiation outside a cylinder of radius ρ_{max} is negligibly small) is split into elementary cells (i, j). Then, to each of the photons, we arbitrarily assign the coordinates at the entrance to the material, (R, Θ), and the free path in the layer, l_d:

$$R = RND(R_0), \quad \Theta = RND(2\pi),$$
$$l_d = -\frac{\ln(1 - RND)}{\mu_d}, \quad (6)$$

where *RND* is a function of random variable.

Upon initialization, a photon is assigned the initial weight $W = h\nu$ at the entrance to the biomaterial. The direction of photon motion is defined by the angles Θ and φ (the former is reckoned from the positive direction of the X axis in the plane XY; the latter, from the positive direction of the Z axis). At the point l_d, an elementary event of absorption takes place; i.e., a part $\Delta W = (\mu_a/\mu_s)W_0$ of the photon energy is absorbed at this point and is added to the energy that has been already absorbed in the elementary volume around this point. At the same point, an elementary event of scattering occurs; that is, the photon acquires a new direction of motion. Accordingly, its weight diminishes, becoming $W_d = W_0 - \Delta W$.

At the point of scattering, a new free path to the point of next scattering is arbitrarily (while with regard to the optical properties of the medium) selected. If the elementary volume into which the photon falls exceeds the biomaterial thickness, one determines the photon reflection and refraction probabilities at the interface by the Frenkel formulas; in other words, the probability that the photon passes to the nth layer from the mth layer is found. The random walk of the photon is traced until it is beyond the domain considered. Then, we solve the nonstationary boundary-value problem of thermal conduction with the method of adaptive finite elements, taking into account the energy absorbed in each of the elementary volumes and assuming that the energy is completely converted to heat.

In general, the simulation of thermal processes in a biological tissue exploits the solution of the three-dimensional Poisson equation

$$\mathrm{div}(\lambda \mathrm{grad} T(x, y, z)) = Q(x, y, z) \quad (7)$$

with boundary conditions including tissue–environment interaction. Here, λ is the thermal conductivity, $T(x, y, z)$ is the desired temperature distribution, and $Q(x, y, z)$ is the volume density of the thermal power in the material due to the absorption of the laser radiation (that is found by solving the optical problem).

Laser beams are usually symmetric about their axis; therefore, the three-dimensional problem can be reduced to two-dimensional. With regard to the arrangement of the tissue layers and the symmetry of the boundary conditions, the computer model can be constructed in the cylindrical coordinate system (r, z) [6]. In our case, the axis of symmetry of the biological tissue irradiated coincides with the optical axis of the laser beam.

In constructing a model for tissue hyperthermia, the proper choice of the boundary conditions is as important as finding the thermal load distribution. If R_0 and R_1 are the radii of the laser beam and of the cylindrical domain of interest, respectively, the boundary condition

$$\left.\frac{\partial T}{\partial r}\right|_{r = P_1, z \in [0, Z_1]} = 0, \quad (8)$$

can then be specified on the surface $r = R_1$ for $R_1 \gg R_0$ (here, Z_1 is the longitudinal size of the domain for which the calculations are performed, or the thickness of the multilayer model).

The value of Z_1 is the depth below which the tissue temperature is stabilized at a certain level. As follows from many experiments, this level is 37°C for a healthy man and Z_1 is roughly equal to 450 μm. Therefore, the boundary condition of the first kind

$$T|_{z = Z_1, r \in [0, R_1]} = T_1 \quad (9)$$

is specified on the surface $z = Z_1$.

In the linearized form, the interaction of the biotissue surface with the environment (via convection or radiation) can be described by the boundary conditions of the third kind

$$\left.\left(\lambda\frac{\partial T}{\partial z} - A(T - T_0)\right)\right|_{z = 0, r \in [0, R_1]} = 0, \quad (10)$$

where λ is the thermal conductivity of the top layer of the biotissue, A is the reduced coefficient of heat transfer, and T_0 is the temperature of the environment where the heat is removed.

At the interface between the ith and the $(i + 1)$th layers, $z = Z_i$, the continuity conditions for the heat flux and temperature are met:

$$\left[\lambda_i \frac{\partial T_i}{\partial z} - \lambda_{i+1} \frac{\partial T_{i+1}}{\partial z}\right]\Bigg|_{Z = Z_i,\, r \in [0, R_1]} = 0, \qquad (11)$$

$$[T_i - T_{i+1}]\big|_{Z = Z_i,\, r \in [0, R_1]} = 0. \qquad (12)$$

The domain for which the calculations are performed consists of subdomains (tissue layers) with greatly differing sizes. The thicknesses of the layers vary from 15 μm for the basic epiderma layer to 565 μm for the upper and lower derma. Moreover, thermal loads are highly localized along both the radial ($0 < r < 1000$ μm) and the longitudinal ($0 < z < 750$ μm) coordinates because of laser radiation absorption. Therefore, the denseness of the nodes of the discrete model grid must substantially be increased to provide an adequate accuracy of the numerical solution [7].

The most effective way to do this is to apply the finite element technique [8]. In this case, the variational problem of extremum of the functional

$$\Phi(T) = \int_V [\lambda(\mathrm{grad}T)^2 - 2QT]dV - 2\int_{S_a} A(T - T_0)^2 dS \qquad (13)$$

is solved.

The property [8] of functional (13), which is related to the process of thermal conduction, is that any minimizing function that acquires the value specified by condition (9) on S_T satisfies both differential equation (7) and boundary conditions (8)–(12).

The continuous variational problem becomes discrete if the domain of interest is split into a finite number of axisymmetric toroidal elements with a triangular meridional cross section. Neighboring elements are assumed to be in continuous contact. Temperature values at the nodes are the basic unknowns. For each of the elements, the function T is defined in the form

$$T = [N^e]\{T\}, \qquad (14)$$

where $[N^e]$ is the matrix specifying the dependence of the shape functions [8] on the coordinates r and z and the column $\{T\}$ contains the nodal values of the function T.

To minimize the functional in all parameters $\{T\}$ over the entire domain, one must solve the set of equations

$$\frac{\partial \Phi}{\partial \{T\}} = 0. \qquad (15)$$

To improve the accuracy of the solution and simultaneously save the computer time, an algorithm for constructing adaptive finite elements has been devised [9, 10].

RESULTS AND DISCUSSION

Using this algorithm, we simulated the propagation of low-intensity Gaussian beams in multilayer biomaterials like integument (Figs. 1–3). Figures 4 and 5 depict the temperature distribution across the depth z at the center of the laser spot and the radial temperature fields on the surface.

From Figs. 1–3, one can see that the absorption and scattering at the integument surface are much more appreciable than in the interior and far exceed the volume-averaged values. Therefore, in the interior, the intensity is much lower than that predicted without considering scattering. In addition, taking into account that the laser beam has a finite size also considerably attenuates the intensity in deep-seated layers.

If the beam has a diameter of less than 1 mm, it strongly diverges in the biomaterial. Most of the energy is released in a relatively thin surface layer ($(l < l_d)$;

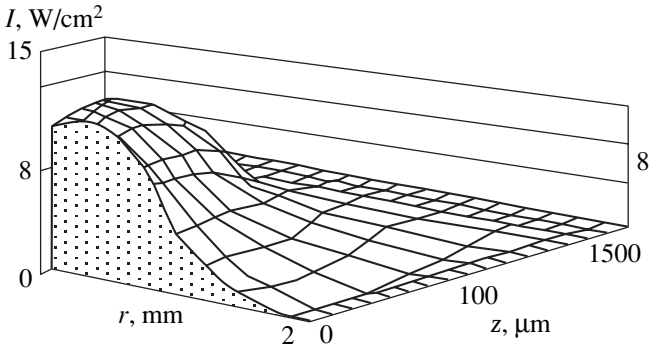


Fig. 1. Intensity distribution in integument for a normally incident Gaussian beam (λ = 633 nm, r_l = 0.5 mm, and P = 70 mW; z is the coordinate along the tissue thickness, and r is the distance to the beam axis).

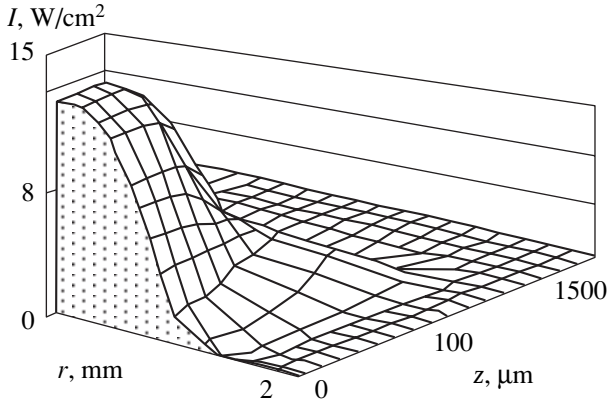


Fig. 2. The same as in Fig. 1 in epiderma.

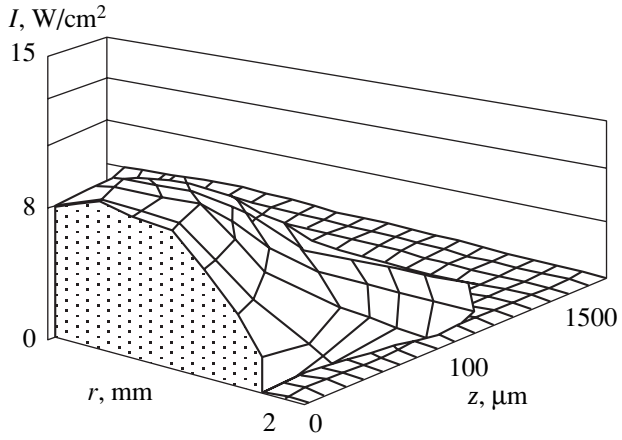


Fig. 3. The same as in Fig. 1 in hypodermic fat.

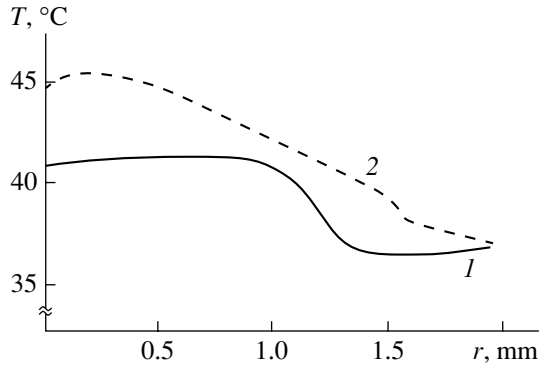


Fig. 4. Temperature distribution over the biotissue surface. (*1*) heavily absorbing tissue; (*2*) heavily scattering tissue. $\lambda = 633$ nm, $\varepsilon_V = 5.3$ J/m^3.

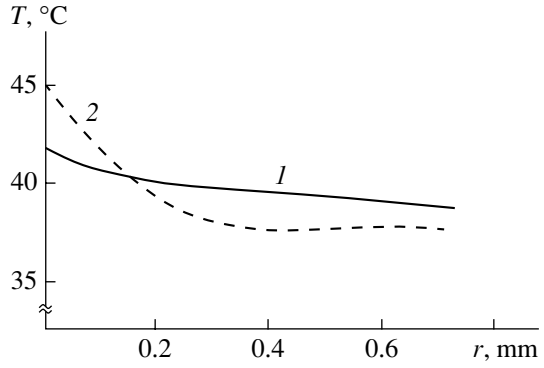


Fig. 5. The same as in Fig. 4 across the thickness.

however, the "tail" of the energy distribution penetrates deep into the surrounding tissues (Fig. 3). To decrease the divergence in a multilayer material and, hence, to raise the optical power delivered to inner layers and organs, beams with a diameter comparable to the photon free path ($l \sim l_d$) are needed. Thus, we are led to the following important conclusion: When blood in the vessels is transdermically irradiated under the conditions of tissue hyperthermia, the power density and the absorbed dose will be high if beams normally incident on the surface are wide and the laser output is high (≈1 W).

Now let us consider temperature fields in multilayer materials. Such information may help in establishing the tissue necrosis area and, hence, in determining the power of lasers suitable for low-intensity therapy.

In heavily scattering biomaterials, the surface temperature is higher than in materials with a low scattering capacity (Fig. 4). The temperature gradient in the former case is also higher, because the temperature drops with depth more rapidly. This means that laser-initiated processes will start first on the integument surface and that the depth of irradiation is smaller than in the case including only absorption. It is important that the beam power must be increased even if the interior is probed by the radiation from the therapeutic window. In this situation, however, the temperature in the surface layer inevitably increases, which may cause tissue destruction and even necrosis. Thus, the application of wide beams remains to be the only way of acting on the inner layers, since the temperature gradients are acceptable for therapy in this case. We emphasize that the necrosis isotherm ($T > 330$ K) for purely absorbing materials is narrower and deeper than that for purely scattering materials, all other things being equal.

The temperature distribution across the thickness is also of great interest. For strongly scattering biotissues, the surface temperature is higher but drops with depth faster. This is of particular importance for laser therapy, since tissue hyperthermia attendant on photodynamic therapy leads to the synergy effect. At the same time, the temperature field on the surface of a weakly scattering tissue is more uniform than when the material scatters intensely. For the latter case, the temperature under the laser spot center is higher presumably because of the action of photons scattered from the beam edges.

These findings support the previous conclusion that the tendency to appreciably irradiate the inner layers in strongly scattering tissues can hardly be justified in laser therapy, since the surface layers may be subjected to power densities and doses causing necrosis.

CONCLUSION

Thus, when calculating the temperature fields due to the laser irradiation of biomaterials, one must take into account both the absorptive and scattering properties of the medium. The surface temperatures of media with equal absorption coefficients but various scattering factors may differ significantly. This circumstance may substantially influence the processes initiated in biotissues.

The temperature fields obtained by the Monte Carlo method agree well with data reported in other papers. This suggests that this method can be applied in more complicated computations. In particular, it may be use-

ful in simulating thermal sources when it is unclear which of models is preferable for solving the thermal conduction equation. Using the method presented, one can predict the sizes of tissue destruction and necrosis areas.

REFERENCES

1. A. V. Priezzhaev, V. V. Tuchin, and L. P. Shubochkin, *Laser Diagnostics in Biology and Medicine* (Nauka, Moscow, 1989).
2. A. S. Kryuk, I. V. Khokhlov, and N. S. Serdyukov, *Therapeutic Effectiveness of Low-Intense Laser Emission* (Nauka i Tekhnika, Minsk, 1986).
3. V. V. Tuchin, Izv. Akad. Nauk, Ser. Fiz. **59** (4), 131 (1993).
4. V. N. Vasil'ev and S. K. Serkov, Inzh.-Fiz. Zh. **63**, 241 (1993).
5. W. F. Cheong, S. A. Prahl, and A. J. Welch, IEEE J. Quantum Electron. **26**, 2166 (1990).
6. A. Sagi, A. Shitzer, A. Katzir, *et al.*, Opt. Eng. **31**, 1417 (1992).
7. R. Graaf, M. H. Koelink, F. F. M. de Mul, *et al.*, Appl. Opt. **32**, 426 (1993).
8. S. V. Patankar, Sovrem. Mashinostr., Ser. A **6**, 1 (1989).
9. Yu. N. Shcherbakov and A. N. Yakunin, Mat. Model. **4** (5), 109 (1990).
10. Yu. N. Shcherbakov and A. N. Yakunin, in *Proceedings of the International Scientific and Technical Conference on Topical Problems in Electron Instrument-Making, Novosibirsk, 1992*, Vol. 1, Part 2, p. 35.

Translated by V. Isaakyan