Optical Properties of Metal Nanoparticles Synthesized in a Polymer by Ion Implantation: A Review

A. L. Stepanov

Kazan Physicotechnical Institute, Russian Academy of Sciences, Kazan, 420029 Tatarstan, Russia
Physics Institute 1, Technical University of Aachen, 52056 Aachen, Germany
e-mail: anstep@kfti.knc.ru
Received December 30, 2002; in final form, July 3, 2003

Abstract—Polymer composite layers irradiated by 30-keV Ag$^+$ ions with doses from 3.1×10^{15} to 7.5×10^{16} cm^{-2} and an ion current of 4 µA/cm^2 are investigated. The composites were examined using Rutherford backscattering (RBS), transmission electron microscopy (TEM), and optical spectroscopy. As follows from electron microscopy and electron microdiffraction data, ion implantation is a promising tool for synthesizing silver nanoparticles in the surface region. The optical density spectra taken of these composites demonstrate that the silver nanoparticles exhibit unusually weak plasma resonance. The formation of silver nanoparticles in layers carbonized by ion implantation is considered. Based on the Mie theory, optical extinction spectra for silver particles in the polymer and carbon matrices are simulated and optical spectra for complex silver core–carbon sheath nanoparticles are calculated. The physics behind the experimental optical spectra of the composite is discussed. © 2004 MAIK "Nauka/Interperiodica".

The problem of designing new polymer-based composite materials containing metal nanoparticles (MNPs) is of current interest particularly in the fabrication of magnetooptic data storages [1], picosecond optical switches, directional connectors, Mach–Zehnder interferometers, etc. [2]. The nonlinear optical properties of these composites stem from the dependence of their refractive index on incident light intensity. This effect is associated with MNPs, which exhibit a high nonlinear susceptibility of the third order when exposed to ultrashort (picosecond or femtosecond) laser pulses [3].

Light-induced electron excitation in MNPs (so-called surface plasma resonance, SPR) [4], which shows up most vividly in the range of linear absorption, gives rise to nonlinear optical effects in the same spectral range. Therefore, in practice, the SPR effect may be enhanced by raising the nanoparticle concentration in the composite, i.e., by increasing the volume fraction of the metal phase (fill factor) in the insulator. Systems with a higher fill factor offer a higher nonlinear cubic susceptibility, all other things being the same [5].

Metal nanoparticles may be embedded in a polymer matrix in a variety of ways. These are chemical synthesis in an organic solvent [4], vacuum deposition on viscous-flow polymers [6], plasma polymerization combined with metal evaporation [7], etc. However, they all suffer from disadvantages, such as a low fill factor or a great spread in size and shape of the particles synthesized, which offsets the good optical properties of composites. One more promising method is ion implantation [8], which provides controllable synthesis of MNPs at various depths under the surface and unlimit-

edly high impurity doses. Work [9] seems to be the pioneering work in this area. By implantation, one can produce almost any metal–insulator (specifically, metal–polymer) composites, as follows from the table, which gives a comprehensive list of references [9–40], MNP shapes, and implantation conditions for various organic matrices. Note that noble metals exhibit the most pronounced SPR effect and, hence, the highest nonlinearity of the MNP optical properties in insulators [4]. However (see, e.g., [33, 34]), even silver nanoparticles, which usually demonstrate intense linear SPR absorption [4], do not show typical SPR spectra in implanted polymers [34, 37]. The aim of this work is therefore to study the SPR-related linear optical properties of MNPs introduced into a polymer matrix by implantation. We compare experimental optical spectra for silver nanoparticles implantation-synthesized in PMMA with model spectra calculated based on the Mie classical electrodynamic theory [4, 41].

1. EXPERIMENTAL

As substrates, we used 1.2-mm-thick PMMA plates, which are optically transparent in a wide spectral range (400–1000 nm) [42]. Ag$^+$ ion implantation (energy 30 keV, doses from 3.1×10^{15} to 7.5×10^{16} cm^{-2}, ion current density 4 µA/cm^2) was performed under a pressure of 10^{-5} Torr at room temperature on an ILU-3 accelerator. In control experiments, PMMA substrates were irradiated by xenon ions with the same process parameters. The samples obtained were examined with TEM and electron microdiffraction (Tesla BM-500 electron

Conditions for MNP synthesis by ion implantation into polyvinylidene fluoride (PVF$_2$), polyimide (PI), polymethyl methacrylate (PMMA), polymethyl methacrylate with phosphorus-containing fragments (PMMA + PF), polyethylene (PE), poly (ethylene terephthalate) (PET), silicone polymer (phenylmethyl-silane resin with tin diethyldicaprilate) (SP), epoxy resin (ER), polycarbonate (PC), and polyetherimide (PEI)

Metal of particles	Material of matrix	Ion energy, keV	Ion dose, cm^{-2}	Ionic current density, μA/cm^2	Matrix temperature, K	Shape of particles	Method of examination	Refs.
Ti	PET	40	2.0×10^{17}	4.5			TEM	[11]
							XRD	
Cr	PET	40	1.0×10^{17}	4.5			TEM	[10]
							XRD	
Fe	PVF$_2$	25	0.1–1.0×10^{17}	–	300		FMR	[9]
Fe	PI	100–150	0.1–1.0×10^{17}	0.1–5		Spherical	TEM	[12]
Fe	PI	40	0.25–1.2×10^{17}	4, 8, 12	300	Spherical particles and their aggregates	TEM	[13]
							FMR	[14]
Fe	PMMA	100–150	0.1–1.0×10^{17}	0.1–5		Spherical	TEM	[12]
Fe	PMMA	40	0.1–6.0×10^{17}	1–10	300	Spherical particles and their aggregates	TEM	[15]
							FMR	[16]
								[17]
								[18]
								[19]
Fe	PMMA + PF	40	1.0–3.0×10^{16}	1–6	300	Same	TEM	[17]
							FMR	
Fe	PE	25	0.1–1.0×10^{17}	–	300		FMR	[9]
Fe	PET	100–150	0.1–1.0×10^{17}	0.1–5		Spherical	TEM	[12]
Fe	PET	40	1.0–3.0×10^{16}	1–6	300		TEM	[17]
							FRM	
Fe	SP	40	0.3–1.8×10^{17}	4	300	Spherical particles and their aggregates	TEM	[20]
								[21]
								[22]
								[23]
Co	PI	100–150	0.1–1.0×10^{17}	0.1–5		Spherical	TEM	[12]
Co	PI	40	0.25–1.2×10^{17}	4, 8, 12	300	Spherical particles and their aggregates	TEM	[14]
Co	SP	40	0.3–1.8×10^{17}	4	300	Same	TEM	[22]
							FMR	[23]
Co	ER	40	0.3–2.5×10^{17}	2–8	300	Spherical, filamentary, tear-shaped, cubic, etc.	TEM	[24]
								[25]
								[26]
								[27]
								[28]
								[29]
								[30]
Cu	PI	150	0.5–1.0×10^{17}	1–5	<360	Spherical	TEM	[31]
Cu	PI	80	5.0×10^{16}	0.1	<630		TEM	[32]
		100						
Cu	PMMA	40	0.1–6.0×10^{17}	1–6	360	″	TEM	[17]
Cu	PMMA + PF	40	1.0–3.0×10^{16}	1–6	360	″	TEM	[17]
Cu	PET	40	0.5–2.0×10^{17}	4–5	360	″	TEM	[10]

Table. (Contd.)

Metal of particles	Material of matrix	Ion energy, keV	Ion dose, cm^{-2}	Ionic current density, $\mu A/cm^2$	Matrix temperature, K	Shape of particles	Method of examination	Refs.
							AFM	[33]
							XRD	
Cu	PET	40	$1.0{-}3.0 \times 10^{16}$	1–6	360	"	TEM	[17]
Zn	PI	150	5.0×10^{17}	1–5	<360	"	TEM	[31]
Pd	PI	100	$0.1{-}1.0 \times 10^{17}$	0.1	<630	"	TEM	[32]
Ag	PI	130	$0.1{-}5.0 \times 10^{17}$	1–3	<630	"	TEM	[32]
Ag	PMMA	30	$1.0{-}7.5 \times 10^{16}$	4	300	Spherical	TEM	[34]
							OS	[35]
								[36]
								[19]
Ag	PET	79	$0.5{-}2.0 \times 10^{17}$	4.5	–	"	TEM	[10]
								[11]
								[37]
Ag	SP	30	$0.6{-}1.8 \times 10^{17}$	4	300	Spherical particles and their aggregates	TEM OS	[20] [21]
								[38]
Ag	ER	30	$0.22{-}7.5 \times 10^{17}$	4	300	Spherical	TEM OS	[39] [2]
								[30]
Pt	PC	106	1.0×10^{17}	–	–	Same	TEM	[40]
Pt	PEI	106	1.0×10^{17}	–	–	"	TEM	[40]

FMR, ferromagnetic resonance; TEM, transmission electron microscopy; AFM, atomic-force microscopy; OS, optical spectroscopy; and XRD, X-ray diffraction.

microscope) and also with RBS (2-MeV ^{4}He$^+$, backscattering angle $\theta = 150°$, van de Graaf accelerator). The resolution in energy was 21 keV or higher, and the ion current density was no more 10 nA. Optical density spectra were recorded with a Hitachi 330 two-beam spectrophotometer in the range 350–800 nm.

Optical spectra of spherical MNPs embedded in various dielectric media were simulated in terms of the Mie electromagnetic theory [41], which allows one to estimate the extinction cross section σ_{ext} for a wave incident on a particle. This value can be related to the light intensity attenuation ΔI after light of initial intensity I_0 passes through a transparent particle-containing dielectric medium. The light absorption and/or scattering by the particles depend on the absorption, σ_{abs}, and/or scattering, σ_{sca}, cross sections with $\sigma_{ext} = \sigma_{abs} + \sigma_{sca}$. According to the Bouguer–Lambert–Beer law [43],

$$\Delta I = I_0(1 - e^{-\#\sigma_{ext}l}), \qquad (1)$$

where l is the optical layer thickness and # characterizes the nanoparticle concentration in the sample. The extinction cross section is proportional to the absorption factor γ: $\gamma = \#\sigma_{ext}$.

Experimental spectral dependences of optical density (OD) are given by

$$OD = -\log(I/I_0) = \gamma \log(e); \qquad (2)$$

hence, for samples with electromagnetically uncoupled particles, we may put $OD \sim \sigma_{ext}$. Therefore, experimental OD spectra are compared with model spectral dependences that are expressed through σ_{ext} found from the Mie theory.

2. ION SYNTHESIS OF METAL NANOPARTICLES

Ion implantation is an effective tool for introducing single impurities into the surface layer to a depth of several micrometers [8]. The surface modification of the material depends on its properties, as well as on ion implantation parameters (ion type and energy, ion current density, target temperature, etc.). A critical implantation parameter is ion dose F_0, which determines the implant amount. Depending on the modification of an
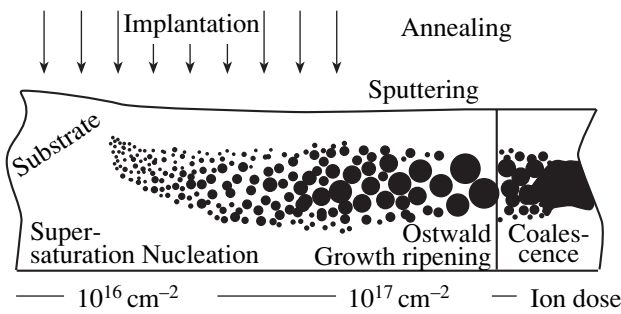
Fig. 1. Basic physical stages of nanoparticle synthesis by ion implantation vs. ion dose.
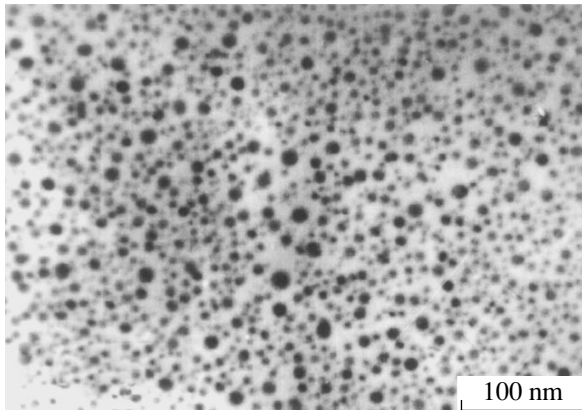

Fig. 2. Micrograph of silver nanoparticles produced by Ag$^+$ implantation into PMMA at a dose of 5×10^{16} cm^{-2}.

insulating target (polymers, inorganic materials, ionic crystals, minerals, etc.), ion implantation may be conventionally divided into low-dose and high-dose implantation (Fig. 1). In the former case ($F_0 \leq 5 \times 10^{14}$ cm^{-2}), the stopped ions are disperse (isolated from one another) in the insulating matrix. The energy of ions implanted is transferred to the matrix through the excitation of electronic shells (ionization) and nuclear collisions. This causes radiation-induced defects, which, in turn, may reversibly or irreversibly modify the material structure [8]. Various types of polymer structure damage have been observed [44]: breaking of covalent bonds in macromolecules, generation of free radicals, cross linkage, oxidation and carbonization of layers irradiated, formation of new chemical bonds between atoms of the insulator or between ions implanted, etc. In addition, ion implantation may be accompanied by the intense sputtering of the surface exposed [37, 45] or, sometimes, by the swelling (ripening) of the polymer [15].

High-dose implantation may also be divided into dose (or time) stages (Fig. 1). At F_0 between 10^{15} and 10^{16} cm^{-2}, the equilibrium solubility of metallic implants in insulators (in particular, polymers) is usu-

ally exceeded, causing the nucleation and growth of MNPs. The dose threshold value depends on the type of the insulator and implant. For 25-keV silver ions implanted into LiNbO$_3$, the threshold dose was found to be $F_0 \approx 5.0 \times 10^{15}$ cm^{-2} [46]; for 30-keV silver ions implanted into epoxy resin, $F_0 \approx 10^{16}$ cm^{-2} [39].

At the next stage of high-dose implantation, starting from $F_0 \geq 10^{17}$ cm^{-2}, the existing MNPs coalesce to form MNP aggregates or quasi-continuous films in the surface layer (Fig. 1). For example, the irradiation of epoxy resin by 40-keV cobalt ions at higher-than-threshold doses favors the formation of thin labyrinth structures [26]. The MNP distribution established in the insulator after coalescence or Ostwald ripening may be disturbed by postimplantation thermal or laser annealing.

In this work, we study composites where MNPs are disperse and isolated from one another, i.e., synthesized at ion doses of 10^{15}–10^{17} cm^{-2}. In our case of implantation by heavy but relatively low (30 keV) Ag$^+$ ions, nuclear collisions prevail in ion–insulator interaction. They displace atoms in the polymer matrix and break some of the chemical bonds in it. Along with this, target atoms effectively lose electrons and the implanted Ag$^+$ ions deionize with the formation of neutral Ag atoms (Ag^0). Basically, Ag atoms may combine with arising organic radicals and polymer ions or take part in the oxidation reaction. However, because of the great difference in Gibbs free energy between Ag atoms and atoms of PMMA elements, Ag–Ag bonding is energetically more favorable.

The formation of MNPs proceeds in several stages: the accumulation of and subsequent supersaturation by Ag^0 atoms in a local surface region of the polymer, the formation of nuclei consisting of several atoms, and the growth of silver particles from the nuclei. Assuming that the nanoparticles nucleate and grow via the successive attachment of silver atoms (which are neutralized embedded Ag$^+$ ions), one may conclude that this process is governed simultaneously by the diffusion coefficient and local concentration of silver atoms, i.e., depends on the matrix temperature. In this work, ion implantation was performed under identical conditions; specifically, the polymer during irradiation was kept at room temperature.

As follows from electron microscopy data, silver ion implantation under the experimental conditions considered results in the formation of silver nanoparticles. For example, the cross-sectional micrograph in Fig. 2 (a dose of 5.0×10^{16} cm^{-2}) shows dark spherical nanoparticles against the bright field (polymer). The irradiation of PMMA by xenon ions did not result in such patterns. Microdiffraction patterns demonstrate that the nanoparticles have the fcc structure of metallic silver. The patterns consist of thin rings (corresponding to polycrystalline nanoparticles) imposed on wide diffuse faint rings from the amorphous polymer matrix. By compar-

ing the experimental diffraction patterns with standard ASTM data, we can conclude that implantation does not form chemical compounds involving silver ions.

From RBS spectra (Fig. 3), it is seen that the silver implantation depth (i.e., the depth where the nanoparticles are located) is virtually independent of the ion dose. Dependence on dose is observed only for the silver peak intensities with the widths and positions of the lines remaining unchanged. It is known that the implantation depth of an ion depends largely on its energy (accelerating voltage) [8] provided that the chemical constitution of the surface irradiated does not change dramatically [47]. The similarity of the RBS spectra shown in Fig. 3 implies that the arising MNPs do not restrict the penetration depth of silver ions at the higher dose. Thus, in PMMA, the implantation dose, being responsible for the amount of the implant, directly influences the MNP size but does not affect the implant distribution profile. As was noted above, the particles nucleate at a dose of $\sim 10^{16}$ cm^{-2} (low-dose implantation). For silver in PMMA, this dose provides MNPs with a size of about 2 nm [47]. However, at a dose of 5.0×10^{16} cm^{-2} (Fig. 2), the particles grow to 10 nm. For 30-keV silver ions implanted into PMMA-like materials, numerical analysis using the TRIM [8] and DYNA [47] statistical algorithms estimates the penetration depth of the ions (the depth of location of the particles) as ≈ 25 nm, which is comparable to the RBS data in order of magnitude.

Experimental optical absorption spectra for PMMA irradiated by xenon and silver ions at various doses are shown in Fig. 4. It is evident that the xenon irradiation of PMMA does not produce nanoparticles, as also follows from the micrographs (Fig. 2). In Fig. 4a, as the xenon ion dose increases, the absorptivity of the polymer in the visible (especially in the close-to-UV) range also increases monotonically. This indicates the presence of radiation-induced structure defects in the PMMA. The absence of absorption bands in these spectral curves is noteworthy. The implantation of silver not only generates radiation-induced defects but also causes the nucleation and growth of nanoparticles. Therefore, along with the absorption intensity variation as in Fig. 4a, an absorption band associated with silver nanoparticles is observed (Fig. 4b). For the lowest dose, the peak of this band is near 420 nm and shifts to longer waves (up to 600 nm) with dose, with the band broadening significantly. The peak of this band is not high, although it is related to the SPR effect in the silver nanoparticles. Such a low intensity of SPR absorption is untypical of silver nanoparticles in PMMA and cannot be explained by the polymer environment of the particles. When silver particles were synthesized in PMMA by the convection technique [48], the SPR intensity was very intense, unlike our experiment. Nor can the weak SPR absorption be explained by any features of the implantation process. For comparison, Fig. 4b shows the optical density spectrum for inor-

Fig. 3. RBS spectra from PMMA irradiated by silver ions for doses of (*1*) 7.7×10^{16} and (*2*) 6.25×10^{15} cm^{-2}.
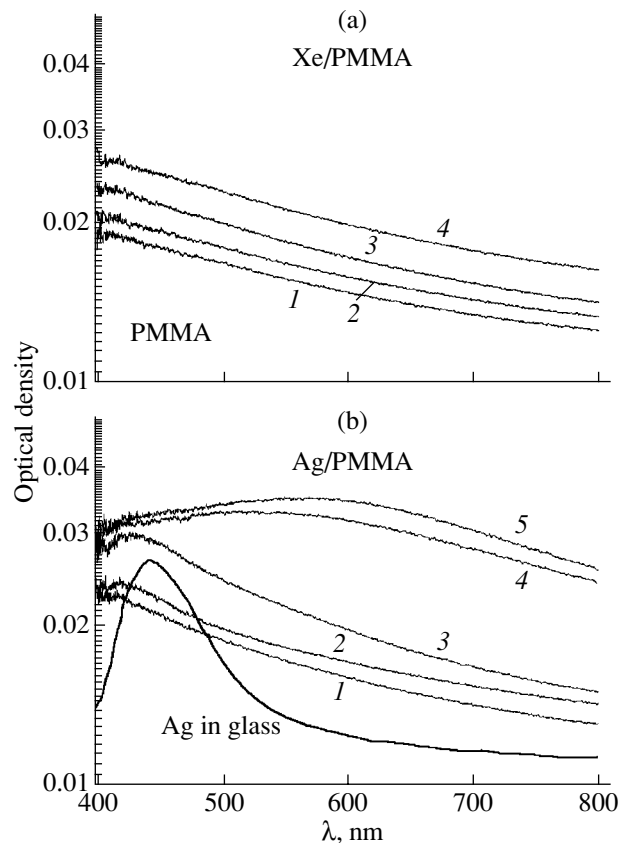
Fig. 4. Optical density spectra from PMMA irradiated by (a) xenon and (b) silver ions for doses of (*1*) 0.3×10^{16}, (*2*) 0.6×10^{16}, (*3*) 2.5×10^{16}, (*4*) 5.0×10^{16}, and (*5*) 7.5×10^{16} cm^{-2}. The spectrum taken from silica glass irradiated by silver ions (5.0×10^{16} cm^{-2}) [48] is shown for comparison.

ganic silica glass (SiO$_2$) irradiated by silver ions under conditions similar to the ion synthesis conditions used in this work (silica glass has refractive index $n \approx 1.5$, close to that of PMMA) [47]. It is seen that the absorp-
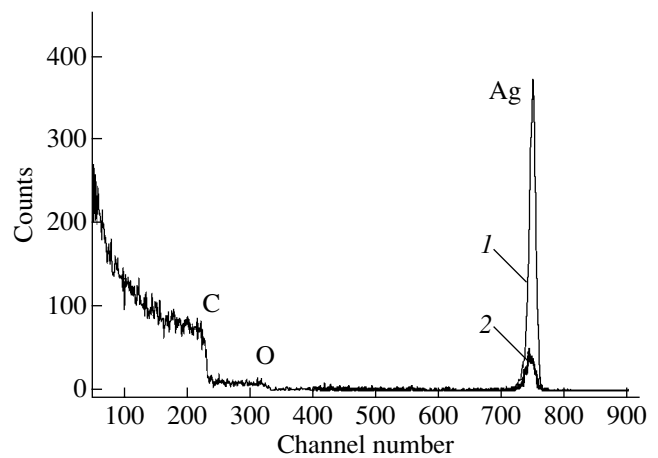
tion of silver nanoparticles in the glass (Fig. 4b) is much more intense (even in view of the background absorption due to matrix structure imperfections) than the absorption of the particles in the polymer. Note that the particle size distributions in the glass and PMMA are nearly the same. Below, we simulate the optical properties of the Ag–PMMA composite and discuss various effects that may clarify the SPR absorption of silver nanoparticles synthesized in PMMA by ion implantation.

3. SIMULATION OF THE OPTICAL PROPERTIES OF COMPOSITES

3.1. Optical Extinction of Silver Nanoparticles in Accordance with the Surrounding Matrix

The attenuation (extinction) of an optical wave propagating in a medium with MNPs depends on the amount of the SPR effect and the light scattering efficiency. The wavelength of optical radiation, the particle size, and the properties of the environment are governing factors in this process. Within the framework of classical electrodynamics (the Maxwell equations), the problem of interaction between a plane electromagnetic wave and a single spherical particle was exactly solved by Mie in terms of optical constants of the interacting objects [41, 49]. According to the Mie theory, the extinction and scattering cross sections are expressed as an infinite sum of spherically symmetric partial electric and magnetic waves that generate fields similar to those generated by the particle when it is viewed as an excited multipole. The extinction cross section is generally given by

$$\sigma_{\text{ext}} = \frac{2\pi}{|k^2|} \sum_{L=1}^{\infty} (2L+1)\text{Re}(a_L + b_L), \qquad (3)$$
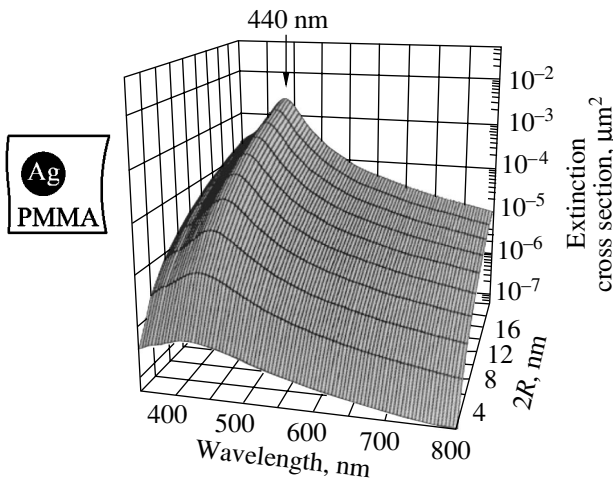
440 nm

Fig. 5. Analytical optical extinction spectra for silver nano-particles embedded in PMMA vs. particle size.

where k is the wavenumber and L is the order of spherical multipole excitation in the particle.

The case $L = 1$ corresponds to a dipole; $L = 2$, to a quadrupole; and so on. The Mie coefficients a_L and b_L are expressed through the Riccati–Bessel cylindrical functions Ψ_L and η_L of variables mx or x, where $mx = \varepsilon_{\text{Ag}}/\varepsilon_{\text{PMMA}}$ is the ratio of the optical constants of the particles and polymer environment and $x = |k|R$ is the dimensional parameter (R is the radius of the particle). These coefficients are given by [49]

$$a_L = \frac{m\Psi_L(mx)\Psi_L'(x) - \Psi_L'(mx)\Psi_L(x)}{m\Psi_L(mx)\eta_L' - \Psi_L'(mx)\eta_L(x)}, \qquad (4)$$

$$b_L = \frac{\Psi_L(mx)\Psi_L'(x) - m\Psi_L'(mx)\Psi_L(x)}{\Psi_L(mx)\eta_L' - m\Psi_L'(mx)\eta_L(x)}. \qquad (5)$$

In the general case, the Mie electromagnetic theory imposes no limitations on the wavelength of optical radiation. Therefore, the operation on the optical constants of the particles and matrix results in extinction spectra, so-called Mie resonance bands [4], which agree well with experiment. However, the Mie theory, which relies on the spectral dependence of the optical constants, does not allow one to penetrate deep into the physics of the Mie optical peaks exhibited by the particles. Yet independent investigations [4] into the behavior of silver nanoparticles showed that Mie resonances are due to the SPR effect, so analytical Mie spectra may be compared with experimental data.

Let us apply Eqs. (3)–(5) from the Mie theory to simulate extinction spectra for silver nanoparticles embedded in a polymer matrix and compare the resulting spectra with the experimental data shown in Fig. 4. In theoretical calculations, we used the complex value of the optical constant ε_{Ag} in the visible range [50] that was obtained by measurements on a set of fine silver particles. Such an approach [50] takes into account limitations imposed on the electron free path in particles of different size and electron scattering at the particle–insulator [52–54] interface and thus yields a more exact value of ε_{Ag} than the procedure of correcting optical constants for bulk silver [51]. The complex values of $\varepsilon_{\text{PMMA}}$ for the polymer matrix were found elsewhere [55]. The extinction was calculated for particles between 1 and 10 nm in size (according to the MNP sizes in Fig. 2).

At the early stage of simulation, consider the simple case where silver nanoparticles are incorporated into the PMMA matrix. Associated extinction spectra for different metal particle sizes are shown in Fig. 5. These spectra feature a wide band, which covers the entire spectral range. In the given range of particle sizes, the position of the SPR absorption maximum (near 440 nm) is independent of the particle size. However, the extinction band intensity grows while the band itself somewhat narrows with increasing particle size. Comparing the analytical and experimental spectra, we see

that, Fig. 5 corresponds to the situation where PMMA is irradiated by silver ions with doses between 0.3×10^{16} and 2.5×10^{16} cm^{-2} (Fig. 4b, curves *1–3*). This dose range corresponds to the early stage of MNP nucleation and growth in the OD spectral band with a maximum between 420 and 440 nm. Thus, one may conclude that ion implantation in this dose range results in the formation of silver nanoparticles, as is also revealed microscopically. It may be supposed that radiation-induced defects in the PMMA have an insignificant effect on the MNP optical properties in this case. However, at higher implantation doses, the recorded OD spectra and the analytical spectra shown in Fig. 5 diverge; hence, the structure of the composite must be refined.

To explain the experimental dependences corresponding to high-dose silver implantation into PMMA, we will first elucidate the difference between implantation into polymers and inorganic insulators (silicate glasses, single crystals, minerals, etc.). The most important distinction is that, as the absorbed dose grows, so does the number of dangling chemical bonds along the track of an ion. Because of this, gaseous hydrogen, low-molecular hydrocarbons (e.g., acetylene), CO, and CO_2 evolve from the matrix [56]. In particular, ion-irradiated PMMA loses $HCOOCH_3$ methoxy groups [57]. The evolution of several organic fractions leads to the accumulation of carbon in the polymer layer irradiated, and radiation-induced chemical processes may cause chain linking. Eventually, an amorphous hydrogenated carbon layer is produced. Polymer carbonization starts with the formation of polycyclic compounds (in essence, primary carboniferous clusters) and, at higher doses, ends up with the formation of the well-developed carbonized phase via carbon cluster linking.

In view of the specific phase structure of the polymer irradiated, it is of interest to study the optical properties (extinction) of silver nanoparticles embedded in the amorphous carbon matrix (a : C-matrix). For this system, the extinction cross section spectra vs. particle size dependence (Fig. 6) was simulated in the same way as for the MNP–PMMA system, i.e., by using complex optical constants ε_C for amorphous carbon, which were taken from [58]. As before (Fig. 4b), throughout the particle size interval, the extinction spectra exhibit a single broad band, which covers the visible range, with a peak at longer waves (510 nm). The calculated longwave shift of the peak, which is observed upon changing the matrix, may be assigned to a longer wave OD band in the experimental spectra for the PMMA, which arises when the silver ion dose exceeds 2.5×10^{16} cm^{-2} (Fig. 4b; curves *3, 4*). It seems that this spectral shift may be associated with the fact that the pure polymeric environment of the silver nanoparticles turns into the
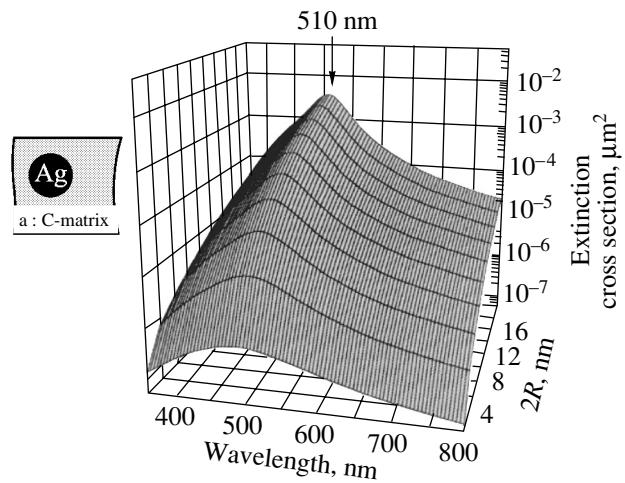

Fig. 6. Analytical optical extinction spectra for silver nanoparticles embedded in the a : C-matrix vs. particle size.

amorphous carbon as the implantation dose rises. The broader extinction bands in the a : C-matrix (Fig. 6) compared with the PMMA (Fig. 5) also count in favor of this supposition, since the broadening of the extinction bands is observed in the experiments as well (Fig. 4b). In a number of experiments [56], however, the carbonization of the polymer surface layer depended on the type of polymer and ion, as well as on the process parameters, and comes to an end at doses of $(0.5–5.0) \times 10^{16}$ cm^{-2}, but the entire material was not carbonized. The carbon clusters may reach several tens of nanometers in size [59]. Thus, the assumption that the polymer irradiated is completely carbonized, which was used in the simulation (Fig. 6), does not meet the real situation when the process lasts for a long time. Below, the variation of the extinction spectra with the amount of carbon in the PMMA layer is analyzed in terms of a model that considers the optical properties of silver MNPs covered by the amorphous carbon sheath.

3.2. Optical Extinction of the Nanoparticles Represented as Sheathed Cores

Extinction spectra for nanoparticles represented as a silver core covered by a carbon sheath in an insulating matrix (PMMA) will be analyzed in terms of the Mie relationships for sheathed cores [60]. Here, an additional interface for which electrodynamic boundary conditions must be set up arises. Plasmon–polariton modes may be excited in both the core and the sheath. These modes, interacting through the inner interface, are responsible for the resulting extinction spectrum. The Mie coefficients a_L and b_L for a homogeneous sphere are replaced by associated expressions for a single-sheath spherical core [61, 62]:

$$a_L = -\frac{m\Psi_L(mx)[\Psi'_L(x) + T_L\vartheta'_L(x)] - \Psi'_L(mx)\Psi_L(x)[\Psi_L(x) + T_L\vartheta_L(x)]}{m\xi_L(mx)[\Psi'_L(x) + T_L\vartheta'_L(x)] - \xi'_L(mx)[\Psi_L(x) + T_L\vartheta_L(x)]}, \quad (6)$$

$$b_L = -\frac{\Psi_L(mx)[\Psi'_L(x) + S_L\vartheta'_L(x)] - m\Psi'_L(mx)[\Psi_L(x) + S_L\vartheta_L(x)]}{\xi_L(mx)[\Psi'_L(x) + S_L\vartheta'_L(x)] - m\xi'_L(mx)[\Psi_L(x) + S_L\vartheta_L(x)]}, \tag{7}$$

where the functions T_L and S_L are given by

$$T_L = -\frac{m\Psi_L(mx)[\Psi'_L(x) + \vartheta'_L(x)] - \Psi'_L(mx)[\Psi_L(x) + \vartheta_L(x)]}{m\vartheta_L(mx)[\Psi'_L(x) + \vartheta'_L(x)] - \vartheta'_L(mx)[\Psi_L(x) + \vartheta_L(x)]}, \tag{8}$$

$$S_L = -\frac{\Psi_L(mx)[\Psi'_L(x) + \vartheta'_L(x)] - m\Psi'_L(mx)[\Psi_L(x) + \vartheta_L(x)]}{\vartheta_L(mx)[\Psi'_L(x) + \vartheta'_L(x)] - m\vartheta'_L(mx)[\Psi_L(x) + \vartheta_L(x)]}. \tag{9}$$

In expressions (6)–(9), as in (4) and (5), L, m, and x refer to the Riccati–Bessel functions Ψ_L, ϑ_L, and ξ_L and the dimensional parameter kR depends on R, where R is the radius of the sheathed core.

Optical extinction spectra for a silver nanoparticle with a fixed size of the core (4 nm) and a varying thickness of the carbon sheath (from 0 to 5 nm) are shown in Fig. 7. The maximum of the SPR bands of the particles is seen to shift from 410 nm (uncovered particle, Fig. 5) to approximately 500 nm. Simultaneously, the SPR band intensity decreases, while the UV absorption increases, so that the absorption intensity at 300 nm and a sheath thickness of 5 nm exceeds the SPR absorption of the particles. Both effects (namely, the shift of the SPR band to longer waves and the increased absorption in the near ultraviolet) agree qualitatively with the variation of the experimental optical density spectra (Fig. 4b) when the implantation dose exceeds 2.5×10^{16} cm^{-2}. Thus, our assumption that the increase in the carbonized phase fraction with implantation dose and the variation of the optical density spectra (Fig. 4b) go in parallel is sustained by the simulation of the optical extinction for complex particles (sheathed cores, Fig. 7).
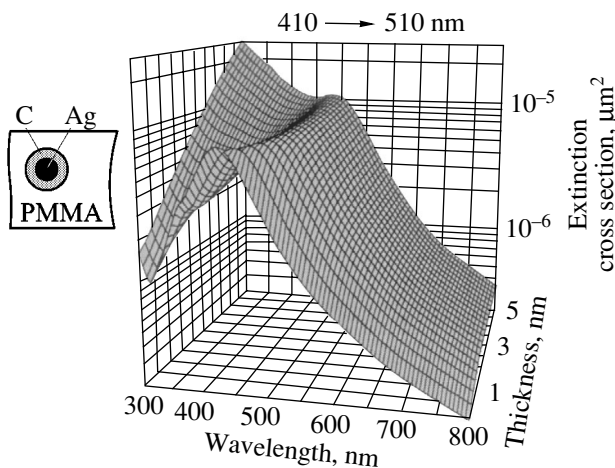

Fig. 7. Analytical optical extinction spectra for 4-nm silver nanoparticles with a carbon sheath that are placed in the PMMA matrix vs. sheath thickness.

In spite of the fact that the model dependences on the carbon sheath thickness and the experimental dose dependences agree qualitatively, discrepancies still exist, particularly, in the position of the long-wave maximum in the optical density spectra and in the breadths of the simulated and experimental spectra. Possible reasons for such quantitative discrepancies are discussed below.

3.3. Effects Arising at the Silver Core–Carbon Sheath Interface and Their Relation to Surface Plasma Resonance Spectra

Interest in carbon-based composites with MNPs goes back a long way. Examples are the studies of magnetic properties of cobalt particles [63], electric and optical properties of layers with copper [64–66] or silver [67, 68] nanoparticles, etc. It was found in optical absorption experiments that copper and silver nanoparticles [64, 68] dispersed in carbon matrices exhibit a weak SPR effect as in our work (Figs. 4b, 6, 7).

When analyzing the optical properties of nanoparticles embedded in a medium, one should take into account effects arising at the particle–matrix interface, such as the static and dynamic redistributions of charges between electronic states in the particles and the environment in view of their chemical constitution [69].

Consider first the static charge redistribution. When an atom is deposited (adsorbed) on the MNP surface, the energy levels of this atom ε_a change their positions compared with those in the free state [69, 70] (Fig. 8). When the number of the adsorbed matrix atoms becomes significant, their contact generates a wide distribution of density of states. Most frequently, the adsorbed atoms are separated from surface atoms of the metal by a tunnel barrier. The gap between the energy positions ε_a of the adsorbed atoms and the Fermi level ε_F of the particles depends on the type of adsorbate (Fig. 8). In addition, the overlap between the energy positions of the matrix atoms and the energy positions of the silver surface atoms depends on the rate with which the electrons tunnel through the barrier. Accordingly, the conduction electron density in the particles embedded will change compared with that in particles placed in a vacuum (without adsorbates): it decreases if

the electrons tunnel toward the adsorbed atoms or increases when the electrons tunnel in the reverse direction. Eventually, equilibrium between the particle and the matrix sets in; i.e., a constant electrical charge (Coulomb barrier) forms at the nanoparticle surface.

Such a static charge redistribution due to the deposition of an adsorbate on the particle surface and the respective change in the electron concentration in the MNPs were also observed in the SPR absorption spectra [4, 69]. In metals (silver, sodium, aluminum, etc.), where free conduction electrons dominate, the SPR spectral maximum $h\omega_{max}$ depends on the concentrations of electrons, N, and nanoparticles as

$$h\omega_{max} \approx [N/(\varepsilon_0 m_{eff})]^{1/2}[2\varepsilon_m + 1 + \chi_1^{inter}]^{1/2}, \quad (10)$$

where ε_m is the permittivity of the matrix, χ_1^{inter} specifies the contribution of the real part of the susceptibility of interband optical transitions in a metal, and m_{eff} is the effective mass of an electron.

It was shown [71] that the incorporation of silver nanoparticles into the carbon matrix of C_{60} fullerene (or the deposition of carbon on the nanoparticle surface) reduces the concentration of $5sp$ electrons in the particle roughly by 20%, since they are trapped by matrix molecules. According to (10), the decrease in N is bound to shift the MNP extinction spectrum toward longer waves, as is also demonstrated by comparing the experimental spectra of the particles in free space (without an adsorbate) with those of the particles in the C_{60} matrix [71]. Samples studied in [71] were similar to those obtained by ion implantation in our work (a carbonized layer near silver particles implanted into the polymer). Thus, the shift of the SPR extinction band into the longer wave range with increasing implantation dose in this experiment (Fig. 4) may be explained by the formation of a carbon sheath around silver nanoparticles. This sheath traps conduction electrons of the particles. The simulation (Fig. 7) also demonstrates the shift of the SPR maximum. However, the effect of dynamic charge redistribution is disregarded in the Mie theory. Therefore, the long-wave shift of the SPR band due to the static charge redistribution at the particle–matrix interface is an additional reason why the experimental spectra are observed at longer waves than the model ones (Figs. 6, 7).

Along with the static charge redistribution at the interface, the charge at the same interface may also change dynamically, i.e., with a high rate [69]. After the static state of the charge has been established and the Fermi level at the interface has been stabilized, the MNP electrons optically excited above the Fermi level (hot electrons) may tunnel (by fluctuations) to the matrix over or through the static barrier (Fig. 8). Levels occupied by the electrons in the intermediate (between the particle and the matrix) state depend on the chemical constitution of the materials. Within a residence lifetime, the electrons may tunnel again from the accep-
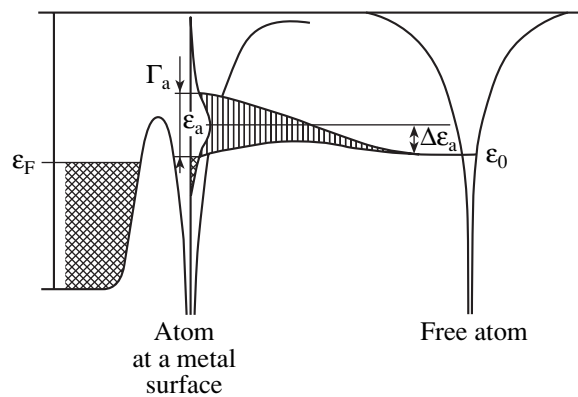

Fig. 8. Electron energy levels in an atom adsorbed on a metal surface [70]. A free atom (to the right) reaches the surface (to the left). Γ_a is the spread of energy levels ε_a. Electron levels in the conduction band of the metal are occupied up to the Fermi level ε_F.

tor levels of the matrix to the particle, and this process may occur over and over again.

The dynamic variation of the charge in time at the particle–matrix interface causes the electron concentration in the particle to fluctuate. This fluctuation directly influences the SPR relaxation. The lifetime of excited conduction electrons in the particle defines the SPR spectral width. Here, the contribution from electron scattering by the interface (because of restrictions imposed on the electron free path [4]) is added to the dynamic variation of the charge at the interface. Thus, the temporal capture of conduction electrons from the particle broadens the SPR-related extinction spectra. This was demonstrated with a set of silver nanoparticles embedded in the C_{60} matrix [71]. Silver nanoparticles in the carbon matrix exhibit a much broader SPR band than in free space. We may therefore suppose that, as the dose rises, the dynamic charge redistribution may broaden the SPR spectra of silver nanoparticles synthesized by ion implantation in PMMA. This is because implantation carbonizes the irradiated layer with increasing absorbed dose and raises the amount of acceptor levels on the MNP surface, which change the relaxation time of the electrons excited. Since the classical Mie theory disregards the dynamic charge redistribution, the model spectra (Fig. 7) must be narrower than the experimental spectra, which is the case.

CONCLUSIONS

In this work, we study the formation of silver nanoparticles in PMMA by ion implantation and optical density spectra associated with the SPR effect in the particles. Ion implantation into polymers carbonizes the surface layer irradiated. Based on the classical Mie electrodynamic theory, optical extinction spectra for silver nanoparticles in the polymeric or carbon environment, as well as for sheathed particles (silver core + carbon sheath) placed in PMMA, as a function of the

implantation dose are simulated. The analytical and experimental spectra are in qualitative agreement. At low doses, simple monoatomic silver particles are produced; at higher doses, sheathed particles appear. The quantitative discrepancy between the experimental spectra and the analytical spectra obtained in terms of the Mie theory is explained by the fact that the Mie theory disregards the static and dynamic charge redistributions at the particle–matrix interface. The influence of the charge redistribution on the experimental optical spectra taken from the silver–polymer composite at high doses, which cause the carbonization of the polymer irradiated, is discussed.

The table, which summarizes available data for ion synthesis of MNPs in a polymeric matrix, and the references cited there may be helpful in practice.

REFERENCES

1. V. V. Randoshkin and A. Ya. Chervonenkis, *Applied Magnetooptics* (Énergoizdat, Moscow, 1990).

2. A. L. Stepanov, R. I. Khaibullin, S. N. Abdullin, and I. B. Khaibullin, RF Patent No. 97,109,708 (1997).

3. S. Muto, T. Kubo, Y. Kurokawa, and K. Suzuki, Thin Solid Films **322**, 233 (1998).

4. U. Kreibig and M. Vollmer, *Optical Properties of Metal Clusters* (Springer, Berlin, 1995).

5. L. Minyung, S. K. Taek, and S. C. Young, J. Non-Cryst. Solids **211**, 143 (1997).

6. S. N. Abdullin, A. L. Stepanov, Yu. N. Osin, and I. B. Khaibullin, Surf. Sci. **395**, L242 (1998).

7. M. Quinten, A. Heilmann, and A. Kiesow, Appl. Phys. B **68**, 707 (1999).

8. P. T. Townsend, P. J. Chandler, and L. Zhang, *Optical Effects of Ion Implantation* (Cambridge Univ. Press, Cambridge, 1994).

9. N. C. Koon, D. Weber, P. Penrsson, and A. I. Shindler, Mater. Res. Soc. Symp. Proc. **27**, 4466 (1984).

10. Y. Wu, T. Zhang, Y. Zhang, *et al.*, Nucl. Instrum. Methods Phys. Res. B **169**, 89 (2000).

11. Y. Wu, T. Zhang, Y. Zhang, *et al.*, Nucl. Instrum. Methods Phys. Res. B **173**, 292 (2001).

12. K. Ogawa, US Patent No. 4,751,100 (1988).

13. R. I. Khaibullin, V. N. Popok, V. V. Bazarov, *et al.*, Nucl. Instrum. Methods Phys. Res. B **191**, 810 (2002).

14. V. N. Popok, R. I. Khaibullin, V. V. Bazarov, *et al.*, Nucl. Instrum. Methods Phys. Res. B **191**, 695 (2002).

15. V. Yu. Petukhov, V. A. Zhikharev, N. R. Khabibullina, and I. B. Khaibullin, Vysokochist. Veshchestva **3**, 45 (1993).

16. V. Yu. Petukhov, V. A. Zhikharev, V. A. Makovskiĭ, *et al.*, Poverkhnost **4**, 24 (1995).

17. V. Yu. Petukhov, M. I. Ibragimova, N. R. Khabibullina, *et al.*, Vysokomol. Soedin., Ser. A **43**, 1973 (2001).

18. V. Petukhov, V. Zhikharev, M. Ibragimova, *et al.*, Solid State Commun. **97**, 361 (1996).

19. V. V. Bazarov, V. Yu. Petukhov, V. A. Zhikharev, and I. B. Khaibullin, Mater. Res. Soc. Symp. Proc. **388**, 417 (1995).

20. R. I. Khaibullin, Yu. N. Osin, A. L. Stepanov, and I. B. Khaibullin, Vacuum **51**, 289 (1998).

21. R. I. Khaibullin, Yu. N. Osin, A. L. Stepanov, and I. B. Khaibullin, Nucl. Instrum. Methods Phys. Res. B **148**, 1023 (1999).

22. R. I. Khaibullin, V. A. Zhikharev, Yu. N. Osin, *et al.*, Nucl. Instrum. Methods Phys. Res. B **166–167**, 897 (2000).

23. B. Z. Rameev, B. Aktas, R. I. Khaibullin, *et al.*, Vacuum **58**, 551 (2000).

24. A. L. Stepanov, R. I. Khaibullin, S. N. Abdullin, *et al.*, Mater. Res. Soc. Symp. Proc. **392**, 161 (1994).

25. A. L. Stepanov, R. I. Khaibullin, S. N. Abdullin, *et al.*, Inst. Phys. Conf. Ser. **147**, 357 (1995).

26. S. N. Abdullin, A. L. Stepanov, R. I. Khaibullin, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **38**, 2574 (1996) [Phys. Solid State **38**, 1412 (1996)].

27. S. N. Abdullin, A. L. Stepanov, R. I. Khaibullin, and I. B. Khaibullin, RF Patent No. 2,096,835 (1996).

28. S. N. Abdullin, A. L. Stepanov, Yu. N. Osin, *et al.*, Surf. Coat. Technol. **106**, 214 (1998).

29. R. I. Khaibullin, S. N. Abdullin, A. L. Stepanov, *et al.*, Pis'ma Zh. Tekh. Fiz. **22** (3), 48 (1996) [Tech. Phys. Lett. **22**, 112 (1996)].

30. I. B. Khaibullin, R. I. Khaibullin, S. N. Abdullin, *et al.*, Nucl. Instrum. Methods Phys. Res. B **127–128**, 685 (1997).

31. K. Yoshida and M. Iwaki, Nucl. Instrum. Methods Phys. Res. B **19–20**, 878 (1987).

32. T. Kobayashi, T. Iwata, Y. Doi, and M. Iwaki, Nucl. Instrum. Methods Phys. Res. B **175–177**, 548 (2001).

33. Y. Wu, T. Zhang, Y. Zhang, *et al.*, Surf. Coat. Technol. **148**, 221 (2001).

34. A. L. Stepanov, S. N. Abdullin, R. I. Khaibullin, *et al.*, Proc. R. Microsc. Soc. **29**, 226 (1994).

35. A. L. Stepanov, S. N. Abdullin, V. Yu. Petukhov, *et al.*, Philos. Mag. B **80**, 23 (2000).

36. A. L. Stepanov, V. N. Popok, I. B. Khaibullin, and U. Kreibig, Nucl. Instrum. Methods Phys. Res. B **191**, 473 (2002).

37. Y. Wu, T. Zhang, A. Liu, *et al.*, Surf. Coat. Technol. **157**, 262 (2002).

38. A. L. Stepanov, R. I. Khaibullin, and I. B. Khaibullin, Philos. Mag. Lett. **77**, 261 (1998).

39. A. L. Stepanov, S. N. Abdullin, R. I. Khaibullin, *et al.*, Mater. Res. Soc. Symp. Proc. **392**, 267 (1995).

40. G. R. Rao, K. Monar, E. H. Lee, and J. R. Treglio, Surf. Coat. Technol. **64**, 69 (1994).

41. G. Mie, Ann. Phys. (Leipzig) **25**, 377 (1908).

42. M. A. Khashan and A. Y. Nassif, Opt. Commun. **188**, 129 (2001).

43. M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, 1969; Nauka, Moscow, 1973).

44. V. B. Odzhaev, I. P. Kozlov, V. N. Popok, and D. V. Sviridov, *Ion Implantation into Polymers* (Belarusk. Gos. Univ., Minsk, 1998).

45. A. L. Stepanov, V. A. Zhikharev, and I. B. Khaibullin, Fiz. Tverd. Tela (St. Petersburg) **43**, 733 (2001) [Phys. Solid State **43**, 766 (2001)].

46. S. Deying, Y. Saito, and S. Suganomata, Jpn. J. Appl. Phys. **33**, L966 (1994).

47. A. L. Stepanov and D. E. Hole, Recent Res. Development Appl. Phys. **5**, 1 (2002).

48. W. Scheunemann and H. Jäger, Z. Phys. **265**, 441 (1973).

49. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983; Mir, Moscow, 1986).

50. M. Quinten, Z. Phys. B **101**, 211 (1996).

51. P. B. Johnson and R. W. Christy, Phys. Rev. B **6**, 4370 (1972).

52. U. Kreibig, J. Phys. F **4**, 999 (1974).

53. A. P. Prishivalko and A. F. Sinyuk, Opt. Spektrosk. **79**, 139 (1995) [Opt. Spectrosc. **79**, 128 (1995)].

54. A. L. Stepanov, Opt. Spektrosk. **91**, 868 (2001) [Opt. Spectrosc. **91**, 815 (2001)].

55. U. Kreibig, G. Bour, A. Hilger, and M. Gartz, Phys. Status Solidi A **175**, 351 (1999).

56. D. V. Sviridov, Usp. Khim. **71**, 1 (2002).

57. B. Pignataro, M. E. Fragala, and O. Puglisi, Nucl. Instrum. Methods Phys. Res. B **131**, 141 (1997).

58. E. D. Palik, *Handbook of Optical Constants of Solids* (Academic, London, 1997).

59. G. R. Rao, Z. L. Wang, and E. H. Lee, J. Mater. Res. **8**, 927 (1993).

60. A. Aden and M. Kerker, J. Appl. Phys. **22**, 1242 (1951).

61. J. Sinzig, U. Radtke, M. Quinten, and U. Kreibig, Z. Phys. D **26**, 242 (1993).

62. J. Sinzig and M. Quinten, Appl. Phys. A **58**, 157 (1994).

63. H. Wang, S. P. Wong, W. Y. Cheung, *et al.*, J. Appl. Phys. **88**, 2063 (2000).

64. V. I. Ivanov-Omskii, A. V. Tolmatchev, and S. G. Yastrebov, Philos. Mag. B **73**, 715 (1996).

65. V. I. Ivanov-Omskii, V. I. Siklitskiĭ, and S. G. Yastrebov, Fiz. Tverd. Tela (St. Petersburg) **40**, 568 (1998) [Phys. Solid State **40**, 524 (1998)].

66. T. N. Vasilevskaya, S. G. Yastrebov, N. S. Andreev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 2088 (1999) [Phys. Solid State **41**, 1918 (1999)].

67. H. Biederman, Z. Chmel, A. Fejfar, *et al.*, Vacuum **40**, 377 (1990).

68. O. Stenzel, H. Kupfer, T. Pfeifer, *et al.*, Opt. Mater. **15**, 159 (2000).

69. U. Kreibig, *Handbook of Optical Properties*, Vol. 2: *Optics of Small Particles, Interfaces, and Surfaces*, Ed. by R. E. Hummel and P. Wissmann (CRC, London, 1997).

70. J. Hölzl, F. Schulte, and H. Wagner, *Solid Surface Physics* (Springer, Berlin, 1979).

71. U. Kreibig, M. Gartz, and A. Hilger, Ber. Bunsenges. Phys. Chem. **101**, 1593 (1997).

Translated by V. Isaakyan

THEORETICAL
AND MATHEMATICAL PHYSICS

On a New Combined Mechanism of Hot Electron Capture in a Semiconductor

Z. S. Kachlishvili, M. G. Khizanishvili, and É. G. Khizanishvili

Dzhavakhishvili State University, Tbilisi, 380028 Georgia
e-mail: Usc@ictsu.tsu.edu.ge
marina@ictsu.tsu.edu.ge
Received June 9, 2003

Abstract—A new combined (thermal–collisional) mechanism of recombination is proposed. The associated capture cross section is calculated. The electric field ranges where this mechanism comes into play and begins to dominate over the Lax cascade mechanism are established. The calculations are performed for *n*-type samples with different concentrations N_0 of neutral impurity atoms and different degrees of compensation *K*. © 2004 MAIK "Nauka/Interperiodica".

It is well known that, under certain conditions, combined mechanisms of charge carrier capture and impurity ionization become dominant over direct processes. An example of the former is the Lax cascade process [1]: an electron is captured on a high-lying level of a trapping center and then drops, passing through quasi-continuously distributed lower lying levels and emitting acoustical phonons. It is common knowledge that the cross section of such a capture is larger than that of Gummel–Lax direct single-phonon capture [2]. Examples of combined mechanisms of impurity atom ionization are photothermal, thermal–field, and electrothermal processes. There are a large number of theoretical and experimental works where combined mechanisms of capture and ionization are discussed (see, for example, [3–9]).

In this paper, we propose a new combined mechanism of recombination the cross section of which, under certain conditions, may be larger than that of Lax giant traps. Its physical idea is the following [10, 11]. In the presence of neutral impurity atoms, a hot electron undergoing inelastic scattering by the atoms loses a part of its energy, exciting them ($1s \longrightarrow 2p$), and is trapped by a positive center. Due to inelastic scattering, the transfer of the electron to a high-lying level becomes much easier and the subsequent transition to the ground state occurs via the cascade mechanism.

Let an impurity semiconductor have a donor concentration N_D and an acceptor concentration N_A. Then, the concentration of positively charged centers is $N_+ = N_A + n$ and the concentration of neutral impurities is $N_0 = N_D - N_A - n$. Here, *n* is the concentration of free electrons in the conduction band.

It is evident that only electrons whose energy ε is higher than Δε (where Δε is the excitation energy) may be responsible for the $1s \longrightarrow 2p$ excitation of neutral centers. In the momentum space of the simple model of

the spectrum, these are the electrons outside the sphere of a radius $\Delta p = \sqrt{2m\Delta\varepsilon}$. The electrons with a momentum *p* that experienced inelastic scattering fall into the small sphere of radius $(p - \Delta p)$. These electrons may (1) undergo quasi-elastic scattering within a time τ and leave this sphere; (2) be accelerated in a field *E* over a time interval $\tau_E = \Delta p/eE$ and acquire the momentum Δp again; and (3) be captured within a time τ_3 by an N_+ attractive center. For certain values of the electric field, concentrations of N_+ and N_0 centers, and lattice temperature, the following inequalities may be fulfilled:

$$\tau_3 < \tau, \quad \tau_E. \tag{1}$$

If this is the case, as soon as the electrons enter into the small sphere, they will be captured on N_+ centers within the time τ_3, having no time to be scattered quasi-elastically or accelerated in the field to the initial state.

Consequently, under conditions (1), the complex inelastic scattering of hot electrons that was considered above (energy loss accompanied by immediate capture) may be considered as a single (composite) process. The lifetime of an electron in the combined process of capture can then be represented as the sum of the characteristic time of the $1s \longrightarrow 2p$ excitation and characteristic time of capture:

$$\tau_{t\text{-}c} = \tau_0 + \tau_+, \tag{2}$$

where $\tau_{t\text{-}c}$ is the lifetime of an electron in the combined capture process, which is below called the thermal–collisional capture mechanism, and τ_0 and τ_+ are the characteristic times of elementary processes in the thermal–collisional mechanism (excitation and capture, respectively).

Based on Eq. (2) and the relationship between the characteristic times and associated probabilities, we

find an expression for the probability of the combined thermal–collisional capture mechanism:

$$W_{t\text{-}c} = \frac{W_0 W_+}{W_0 + W_+}, \qquad (3)$$

where the subscripts have their initial meaning.

As is seen from expression (3), condition (2) yields the reduced probability of the combined process: it is lower than the partial probabilities of the constituent processes. Physically, this is quite clear.

The probability that an electron will be inelastically scattered by a neutral center per unit time (i.e., that the electron will be transferred from the state with an energy ε to the state with an energy $\varepsilon - \Delta\varepsilon$ and the neutral impurity center, from the level $1s$ to the level $2p$) is given by

$$W_0 = N_0 \sigma_0(\varepsilon) v(\varepsilon). \qquad (4)$$

Here, $v(\varepsilon)$ is the velocity of the electron,

$$\sigma_0 = \pi a_0^2 \left(\frac{\text{Ry}}{\Delta\varepsilon}\right)^2 \left(\frac{E_1}{E_0}\right)^{3/2} \frac{\vartheta_{\chi_{\min}}(\alpha_0, \alpha_1)}{2l_0 + 1} \Phi(U) \qquad (5)$$

is the inelastic scattering cross section by a neutral center [12], Ry is the Rydberg unit, $\vartheta_{\chi_{\min}}(\alpha_0, \alpha_1)$ is the angular factor depending on the quantum numbers of the angular momenta of the states α_0 and α_1, l_0 is the orbital quantum number,

$$\Phi(U) = \left(\frac{U}{U+1}\right)\left(\frac{C}{U+\varphi}\right), \qquad (6)$$

and C and φ are the parameters given in [12].

The probability of cascade capture of an electron with an energy $\varepsilon - \Delta\varepsilon$ on a positive center can also be written in form (4):

$$W_+ = N_+ \sigma_+(\varepsilon - \Delta\varepsilon) v(\varepsilon - \Delta\varepsilon). \qquad (7)$$

Here, according to [13],

$$\sigma_+(\xi) = \frac{1}{3}\frac{4^6 \sigma_1}{\gamma^3 \xi (\xi/\gamma + \delta_0)^3}\left(1 - \exp\left(-\frac{\xi/\gamma + \delta_0}{8}\right)\right), \quad (8)$$

where

$$\xi = \frac{2\varepsilon}{mS^2}, \quad \gamma = \frac{2k_0 T}{mS^2},$$

S is the speed of sound in the semiconductor, δ_0 is the minimal binding energy, σ_1 is a factor that has the dimension of the cross section, and m is the electron effective mass.

In view of (4) and (7), expression (3) takes the form

$$W_{t\text{-}c} = \frac{N_0 v(\varepsilon)\sigma_0(\varepsilon)N_+ v(\varepsilon - \Delta\varepsilon)\sigma_+(\varepsilon - \Delta\varepsilon)}{N_0 v(\varepsilon)\sigma_0(\varepsilon) + N_+ v(\varepsilon - \Delta\varepsilon)\sigma_+(\varepsilon - \Delta\varepsilon)}. \quad (9)$$

Now, using expression (9), we find the effective differential cross section of thermal–collision capture.

From (9), it is seen that $W_{t\text{-}c}$ is proportional to $N_0 N_+$. Then, starting from the general relationship between the probability and corresponding differential cross section and keeping the proportionality between $W_{t\text{-}c}$ and $N_0 N_+$ in view of the dimension, one can write

$$W_{t\text{-}c} = \sigma_{t\text{-}c}(\varepsilon) v(\varepsilon) \frac{N_0 N_+}{N_0 + N_+}. \qquad (10)$$

With (9) and (10), the effective differential cross section of the combined thermal–collision capture mechanism is given by

$$\sigma_{t\text{-}c}(\varepsilon) = \frac{\sigma_0(\varepsilon)\sigma_+(\varepsilon - \Delta\varepsilon)\sqrt{1 - \dfrac{\Delta\varepsilon}{\varepsilon}}}{(1-K)\sigma_0(\varepsilon) + K\sqrt{1 - \dfrac{\Delta\varepsilon}{\varepsilon}}\,\sigma_+(\varepsilon - \Delta\varepsilon)} \qquad (11)$$

if $N_A \gg n$. Here, $K \equiv N_A/N_D$ is the degree of compensation of the sample.

From (11), the differential cross section of the combined process depends on the degree of compensation and the ratio between the excitation energy and the energy of a free electron. Consider particular cases.

(1) $K \ll 1$. In this case, the second term in the denominator of Eq. (11) can be omitted irrespective of the ratio between ε and $\Delta\varepsilon$. Then, we come to

$$\sigma_{t\text{-}c}(\varepsilon) = \frac{\sigma_+(\varepsilon - \Delta\varepsilon)\sqrt{1 - \dfrac{\Delta\varepsilon}{\varepsilon}}}{(1-K)} \qquad (11a)$$

$$\approx \sigma_+(\varepsilon - \Delta\varepsilon)\sqrt{1 - \frac{\Delta\varepsilon}{\varepsilon}}.$$

(a) At $\varepsilon \gg \Delta\varepsilon$,

$$\sigma_{t\text{-}c}(\varepsilon) \cong \sigma_+(\varepsilon - \Delta\varepsilon). \qquad (11b)$$

(b) At $\varepsilon \geq \Delta\varepsilon$, $\sigma_{t\text{-}c}$ is given by expression (11a).

(2) $K \leq 1$.

(a) At $\varepsilon \geq \Delta\varepsilon$,

$$\sigma_{t\text{-}c}(\varepsilon) \cong \sigma_0(\varepsilon). \qquad (11c)$$

(b) At $\varepsilon \geq \Delta\varepsilon$, $\sigma_{t\text{-}c}(\varepsilon)$ is given by (11c) again. This is explained by the fact that, in the case considered, the terms in the denominator of (11) contain the quantities $(1-K)$ and $\sqrt{1 - (\Delta\varepsilon/\varepsilon)}$, which are of the same order of smallness, but $\sigma_+(\varepsilon - \Delta\varepsilon)$ has a sharp peak at $\varepsilon \geq \Delta\varepsilon$ (see (8)).

Figure 1 shows the energy dependence of the ratio $\sigma_{t\text{-}c}(\varepsilon)/\sigma_+(\varepsilon)$ (see (11) and (8)) for different degrees of compensation.

For low degrees of compensation (curves 1–3), the cross section of the thermal–collision capture mechanism is governed, as was expected, by the dependence $\sigma_+(\varepsilon - \Delta\varepsilon)$ and is a decreasing function of the energy of a captured electron. As the degree of compensation
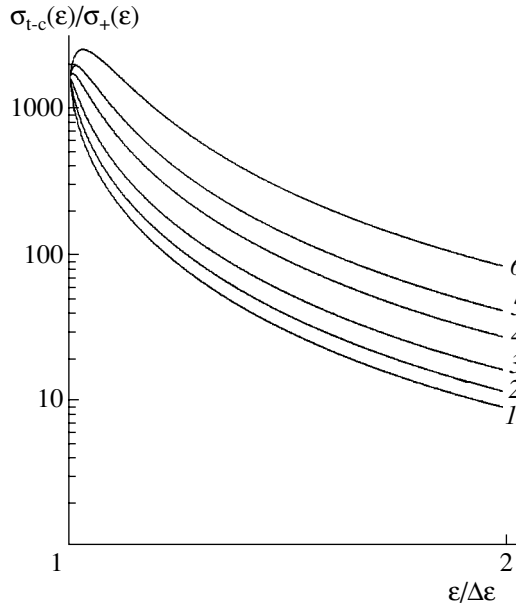
Fig. 1. $\sigma_{t\text{-}c}(\varepsilon)/\sigma_+(\varepsilon)$ on $\varepsilon/\Delta\varepsilon$ for a degree of compensation $K = (1)$ 0.1, (2) 0.3, (3) 0.5, (4) 0.7, (5) 0.8, and (6) 0.9.

Fig. 2. Relationship between the lifetimes (1) τ_3, (2) τ_E, and (3) τ for the sample with $N_0 = 5 \times 10^{15}$ cm^{-3} and $K = 0.5$.

increases, the range where $\sigma_0(\varepsilon)$ controls the run of the dependence $\sigma_{t\text{-}c}(\varepsilon)$ (the rising parts of the curves) widens (curves 4–6).

Now, we find the conditions under which inequalities (1) are valid. From the above reasoning, it follows that the mechanism responsible for momentum dissipation is scattering by neutral and ionized impurity atoms, while the energy is dissipated by acoustical phonons. Under these conditions, the symmetric part of the distribution function takes the form [10]

$$f_0(x) = \exp\left\{-\int^x \frac{dx}{1 + \dfrac{E^2}{E_0^2}\dfrac{1}{\sqrt{x}+\dfrac{l_N^0}{l_I^0}\dfrac{1}{x}}}\right\}, \quad (12)$$

where

$$\frac{1}{E_0^2} = \frac{e^2 l_N^0 l_a^0}{3(k_0 T)^2}; \quad (13)$$

l_I^0 and l_N^0 are the electron-energy-independent parts of the free paths when the momentum is dissipated by impurity ions and neutral centers, respectively; and l_a^0 is the free path that takes into account only energy dissipation by acoustical lattice vibrations.

For calculations, we write (in explicit form) expressions for the times entering into the system of inequal-
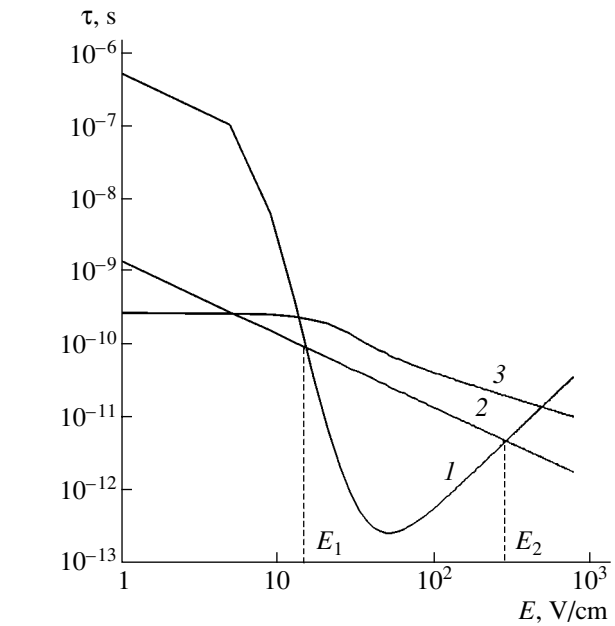
ities (1). The time τ_3 of capturing an electron by an N_+ attractive center (after the electron has been scattered by a neutral impurity atom) and, accordingly, has lost energy $\Delta\varepsilon$ is given by

$$\tau_3 = \frac{1}{N_+ \langle \sigma_+(\varepsilon - \Delta\varepsilon) v_+(\varepsilon - \Delta\varepsilon)\rangle}. \quad (14)$$

Averaging of the recombination rate $\langle \sigma v \rangle$ is performed with distribution function (12) in accordance with the conditions of our problem:

$$\langle \sigma(x - \Delta x) v(x - \Delta x)\rangle$$

$$= \frac{\displaystyle\int_{\Delta x}^{\infty} \sigma(x - \Delta x) v(x - \Delta x)\sqrt{x} f_0(x)dx}{\displaystyle\int_0^{\infty} \sqrt{x} f_0(x)dx}. \quad (15)$$

For the quantities τ_E and τ on the right of inequalities (1), we have

$$\tau_E = \frac{\sqrt{2\Delta x m(k_0 T)}}{eE},$$

$$\tau = \frac{\pi\hbar\rho}{2\sqrt{2}m^{5/2}E_c^2(k_0 T)^{1/2}}(\bar{x})^{-1/2}. \quad (16)$$

The problem is solved graphically. The electric field range $\Delta E \equiv E_2 - E_1$ where the system of inequalities (1) holds is assumed to be bounded by the points of intersection of the curves τ_3 and τ_E, since the inequality $\tau_3 <$
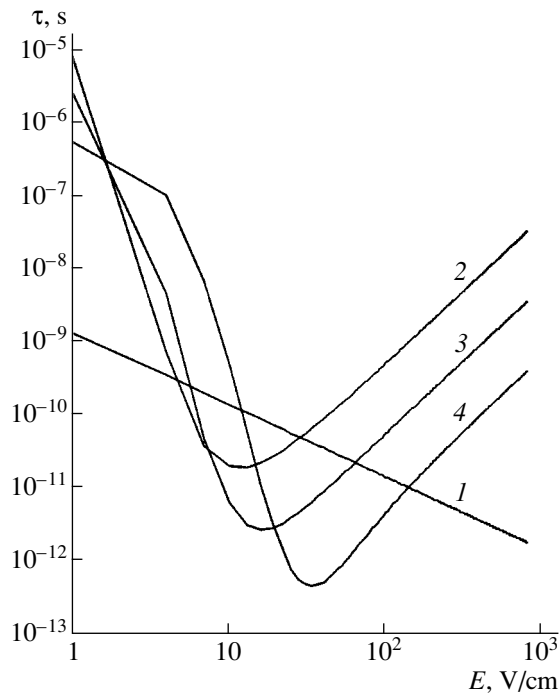
Fig. 3. Lifetimes (*1*) τ_E and (*2–4*) τ_3 for the sample with $N_0 = 5 \times 10^{14}$ cm^{-3} and degree of compensation $K = $ (*2*) 0.1, (*3*) 0.5, and (*4*) 0.9.
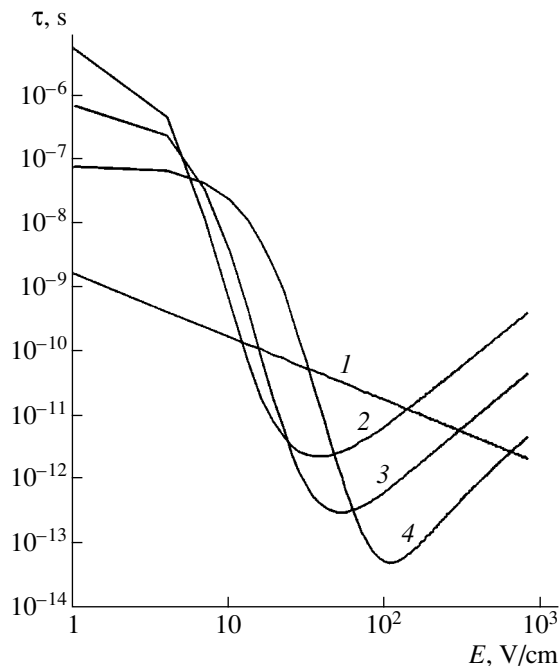

Fig. 4. Lifetimes (*1*) τ_E and (*2–4*) τ_3 for the sample with $N_0 = 5 \times 10^{15}$ cm^{-3}. The degrees of compensation are the same as in Fig. 3.

τ is valid in a wider region of E than the inequality $\tau_3 < \tau_E$ (Fig. 2).

The values of E_1 and E_2 vary significantly with K and N_0. Figures 3 and 4 show the ranges ΔE in which inequalities (1) are valid for $N_0 = 5 \times 10^{14}$ and 5×10^{15} cm^{-3}.

As follows from Figs. 3 and 4, the range where inequalities (1) are valid tends to higher fields with increasing K and N_0. This is not surprising, since the more intense the scattering by impurity ions and neutral centers, the stronger the fields providing a desired mean energy of electrons, in particular, an energy $\Delta\varepsilon$ necessary for inelastic scattering.

To conclude, for a certain low-temperature concentration of neutral and ionized impurity atoms, there invariably exists an applied electric field range where the thermal–collision mechanism of hot electron capture is dominant.

REFERENCES

1. M. Lax, Phys. Rev. **119**, 1502 (1960).
2. H. Gummel and M. Lax, Ann. Phys., No. 2, 28 (1957).
3. V. N. Abakumov and I. N. Yassievich, Zh. Éksp. Teor. Fiz. **71**, 657 (1976) [Sov. Phys. JETP **44**, 345 (1976)].
4. D. V. Berman, T. I. Domasheva, and A. G. Yukov, Fiz. Tekh. Poluprovodn. (Leningrad) **7**, 1882 (1973) [Sov. Phys. Semicond. **7**, 1256 (1973)].
5. V. Karpus and V. I. Perel', Zh. Éksp. Teor. Fiz. **91**, 2319 (1986) [Sov. Phys. JETP **64**, 1376 (1986)].
6. S. V. Kryuchkov and G. A. Syrodoev, Fiz. Tekh. Poluprovodn. (Leningrad) **25**, 655 (1991) [Sov. Phys. Semicond. **25**, 397 (1991)].
7. V. G. Jackeli, Soobshch. Akad. Nauk Gruz. **145**, 525 (1992).
8. P. N. Krylov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 304 (2000) [Semiconductors **34**, 300 (2000)].
9. V. G. Jackeli, Z. S. Kachlishvili, and É. G. Khizanishvili, Izv. Vyssh. Uchebn. Zaved., Tomsk, 1999.
10. Z. S. Kachlishvili, Phys. Status Solidi A **33**, 15 (1976).
11. V. G. Jackeli, Z. S. Kachlishvili, M. G. Khizanishvili, and É. G. Khizanishvili, Soobshch. Akad. Nauk Gruz. **160** (3) (1999).
12. A. A. Vaǐnshteǐn, I. I. Sobel'man, and E. A. Yukov, *Atomic Excitation and Spectral Line Broadening* (Nauka, Moscow, 1979), p. 320.
13. T. O. Gegechkory, V. G. Jackeli, and Z. S. Kachlishvili, Phys. Status Solidi B **112**, 379 (1982).

Translated by Yu. Vishnyakov

Thermal Nonlinearity in a Photoacoustic Cell

U. Madvaliev[1], T. Kh. Salikhov[1, 2], and D. M. Sharifov[1]

[1] *Umarov Physicotechnical Institute, Academy of Sciences of Tajikistan, Dushanbe, 734063 Tajikistan*
e-mail: ymarhon@ac.tajik.net
[2] *Tajik State University, Dushanbe, 734025 Tajikistan*
e-mail: salikhov@ac.tajik.net
Received June 2, 2003

Abstract—The effect of the thermal nonlinearity caused by the temperature dependence of the specific heat and thermal conductivity of a sample, substrate, and air on the temperature field in a photoacoustic cell is studied theoretically. Exact solutions are obtained for a steady temperature field with allowance for thermal nonlinearity and for a nonsteady field without allowance for this nonlinearity. The nonsteady nonlinear problem was solved numerically. It is shown that, due to thermal nonlinearity, the linear dependence of the temperature of the irradiated surface on the heating beam intensity gradually transforms into a power-law dependence as the beam intensity increases. © 2004 MAIK "Nauka/Interperiodica".

INTRODUCTION

Photoacoustic (PA) spectroscopy and related diagnostics are widely used to study condensed media [1–3]. Due to their contactless character, these diagnostics are successfully employed to determine optical, thermophysical, acoustic, and chemical properties of various media when the conventional methods turn out to be inefficient [4–6]. Information provided by these diagnostics makes it possible to examine various inhomogeneous, multicomponent, multilayer, porous, and chemically active systems under real (sometimes extreme) operating conditions [1, 3, 5, 7]. However, this information is usually gained against the background of an established temperature field (TF), and PA measurements are usually performed after the system has passed to this quasi-steady state. The time during which a steady state is established can be roughly estimated as $\tau \sim L^2/\chi$, where L is the scale length and χ is the thermal diffusivity of the medium. The higher the intensity of incident radiation I_o, the higher the temperature of the absorbing sample; as a result, at high heating intensities, the effect of the temperature dependence of the thermophysical and optical parameters of the sample material on the PA signal becomes more pronounced and the estimation of τ is no longer an easy matter.

The necessity of taking into account the effect of the temperature dependence of the thermophysical and optical parameters on the PA signal was first addressed in [8]. However, the specific features of TFs in case of the gas-microphone detection of the PA signal were not considered. Due to the temperature nonlinearity of the thermal conductivity $\kappa(T)$, the second-harmonic PA signal was detected and investigated using the mirage effect [9] and infrared radiometry [10]. In [11], the influence of temperature nonlinearity on the parameters of thermal waves was studied both theoretically and

experimentally and the possibility of determining the nonlinear parameters of some metals was demonstrated. The specific features of the generation of second-harmonic thermal waves were studied theoretically in [12].

In classical studies [13–15] (in [13], the theory of the PA effect was developed), expressions describing a steady TF in a system with constant thermophysical and optical parameters were obtained. However, nonsteady TFs in condensed media in a PA cell were not considered even in the linear approximation. Probably, it is for this reason that, in [16], a nonsteady TF in a one-dimensional sample irradiated by a rectangular laser pulse was calculated numerically.

This study is aimed at a detailed theoretical analysis of the influence of thermal nonlinearities on a TF in highly absorptive and low-conductivity media. We assume that the sample is placed into a gas-microphone cell to perform integral and spectral PA measurements. Thermal nonlinearities are assumed to be related to the temperature dependence of the thermophysical parameters of all three layers within the PA cell. A nonsteady TF is determined in the linear approximation for the same experimental layout.

1. MATHEMATICAL MODEL AND THE PHOTOACOUSTIC CELL LAYOUT

Let us consider a one-dimensional PA cell [13] consisting of three layers: a gas layer (g), a sample (s), and a substrate (b) (Fig. 1). The gas and the substrate are assumed to be transparent for the incident beam with a harmonically modulated intensity. The sample is a highly absorptive medium with a low thermal conductivity and an absorption factor β satisfying the condition $\beta l \gg 1$. In this case, the heat transferred to the

medium is localized within a thin (with a thickness of $\approx \beta^{-1}$) surface layer in which the temperature may rise substantially. In media with a high thermal conductivity (e.g., in metals), heat is rapidly redistributed over the sample and a local equilibrium TF is established. We also assume that the reflection coefficient R and the absorption factor β of the medium remain constant during PA measurements.

We consider the temperature variation range that is far from the phase transition points and the thermal destruction of polymers. We assume that the heat capacity per unit volume, $C(T) = \rho(T)C_p(T)$ (where ρ is the mass density and C_p is the specific heat at a constant pressure), and the thermal conductivities $\kappa_i(T)$ of the layers are linear functions of the temperature:

$$C_i = C_i^{(0)}(1 + \delta_{1i}T'),$$

$$\kappa_i = \kappa_i^{(0)}(1 + \delta_{2i}T'); \quad i = g, s, b, \tag{1}$$

where T' is the temperature increment and

$$\delta_{1i} = \frac{1}{C_i^{(0)}}\frac{\partial C_i}{\partial T}, \quad \delta_{2i} = \frac{1}{\kappa_i^{(0)}}\frac{\partial \kappa_i}{\partial T} \tag{2}$$

are the temperature factors of the thermophysical parameters. Here, the superscript (0) stands for the initial values at T_0 and variations in the system parameters are considered to be small compared to their initial values.

Thus, the set of nonlinear heat conduction equations for a three-layer one-dimensional PA cell takes the form

$$\rho_g(T_g)C_{pg}(T_g)\frac{\partial T_g'}{\partial t} = \frac{\partial}{\partial x}\left(\kappa(T_g)\frac{\partial T_g'}{\partial x}\right), \tag{3}$$

$$0 \le x \le l_g,$$

$$\rho_s(T_s)C_{ps}(T_s)\frac{\partial T_s'}{\partial t} = \frac{\partial}{\partial x}\left(\kappa_s(T_s)\frac{\partial T_s'}{\partial x}\right)$$

$$+ \frac{1}{2}\beta I_0(1-R)\exp(\beta x), \quad -l \le x \le 0, \tag{4}$$

$$\rho_b(T_b)C_{pb}(T_b)\frac{\partial T_b'}{\partial t} = \frac{\partial}{\partial x}\left(\kappa_b(T_b)\frac{\partial T_b'}{\partial x}\right), \tag{5}$$

$$-(l_b + l) \le x \le -l.$$

Since the set of Eqs. (3)–(5) consists of three second-order equations, it should be complemented with six boundary conditions, namely, the continuity of the temperature and heat flux at the cell boundaries, as well as at the gas–sample and sample–substrate interfaces. These conditions are written as

$$T_g'(l_g, t) = T_s'(-l_s - l_b, t) = 0,$$

$$T_g'(0, t) = T_s'(0, t), \quad T_s'(-l_s, t) = T_b'(-l_s, t), \tag{6}$$
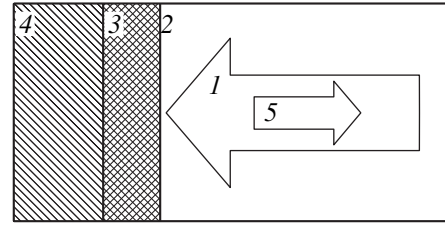
Fig. 1. Geometry of the problem: (*1*) incident amplitude-modulated laser beam, (*2*) buffer gas (air), (*3*) sample, (*4*) substrate, and (*5*) partially reflected beam.

$$\kappa_g(T_g)\frac{\partial T_g'}{\partial x}\bigg|_{x=0} = \kappa_s(T_s)\frac{\partial T_s'}{\partial x}\bigg|_{x=0},$$

$$\kappa_s(T_s)\frac{\partial T_s'}{\partial x}\bigg|_{x=-l} = \kappa_b(T_b)\frac{\partial T_b'}{\partial x}\bigg|_{x=-l}. \tag{7}$$

The initial conditions are trivial:

$$T_g'(0, x) = T_s'(0, x) = T_b'(0, x) = 0. \tag{8}$$

The set of Eqs. (3)–(5), together with boundary and initial conditions (6)–(8), presents a mathematical formulation of the problem of determining a TF in a one-dimensional PA cell. In what follows, we will apply this model to solve the problems formulated above. We note that the problem of thermal waves (i.e., the oscillating part of temperature variations) is beyond the scope of this study.

2. EFFECT OF THERMAL NONLINEARITY ON A STEADY TEMPERATURE FIELD

In a steady state, all the derivatives are zero and the temperature variations are time-independent. In this case, we are dealing only with a spatial distribution of the temperature, and the set of Eqs. (3)–(5) reduces to

$$\frac{d}{dx}\left[\kappa_g(T)\frac{dT_g}{dx}\right] = 0, \quad 0 \le x \le l_g, \tag{9}$$

$$\frac{d}{dx}\left[\kappa_s(T)\frac{dT_s}{dx}\right] = -\frac{1}{2}\beta I_0(1-R)\exp(\beta x), \tag{10}$$

$$-l \le x \le 0,$$

$$\frac{d}{dx}\left[\kappa_b(T)\frac{dT_b}{dx}\right] = 0, \quad -(l + l_b) \le x \le -l. \tag{11}$$

Using the representation $T_i(x) = \delta_{2i}^{-1} g_i(x)$, solutions to Eqs. (9)–(11) with boundary conditions (6) and (7) can be written in the following analytical form:

$$g_g(x) = \left[1 + \Theta_0\delta_{2g}(2 + \Theta_0\delta_{2g})\left(1 - \frac{x}{l_g}\right)\right]^{1/2} - 1, \tag{12}$$
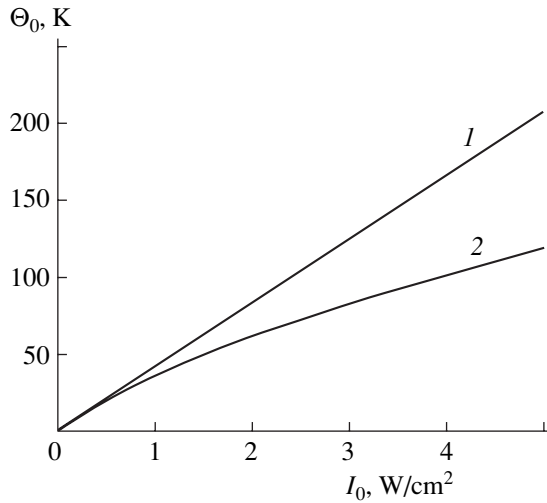
Fig. 2. Temperature of the irradiated surface vs. incident beam intensity: (*1*) linear theory and (*2*) numerical solution of Eqs. (15) and (16).

$$g_s(x) = \left[1 + \delta_{2s}\left(\Theta_0(2 + \delta_{2s}\Theta_0)\left(1 + \frac{x}{l}\right) \right.\right.$$
$$\left.\left. - W_0(2 + \delta_{2s}W_0)\frac{x}{l} \right) + \frac{AI_0}{\beta\kappa_s^{(0)}}\left(1 + \frac{x}{l} - \exp(\beta x) - \frac{x}{l}E\right) \right]^{1/2} - 1, \tag{13}$$

$$g_b(x) = \left[1 + W_0\delta_{2b}(2 + W_0\delta_b)\left(1 + \frac{x+l}{l_b}\right) \right]^{1/2} - 1, \tag{14}$$

where W_0 is the temperature perturbation at the sample–substrate interface, $E = \exp(-\beta l)$, and $A = 1 - R$. The values of Θ_0 and W_0 are determined by numerically solving the set of algebraic equations

$$\Theta_0^2 b_{11} + \Theta_0 b_{12} - W_0 b_{13} - W_0^2 b_{14} + F_1 = 0, \tag{15}$$

$$W_0^2 c_{11} + W_0 c_{12} - \Theta_0 c_{13} - \Theta_0^2 c_{14} - F_2 = 0. \tag{16}$$

Here, the following notation is introduced:

$$b_{11} = \frac{1}{2}(\delta_{2s} + a_1\delta_{2g}), \quad b_{12} = 1 + a_1, \quad a_1 = \frac{l\kappa_g^{(0)}}{l_g\kappa_s^{(0)}},$$

$$b_{13} = c_{13} = 1, \quad b_{14} = \frac{1}{2}\delta_{2s} = c_{14}, \quad c_{12} = 1 + \frac{a_1}{a_2},$$

$$c_{11} = \frac{1}{2}(\delta_{2s} + a_2\delta_{2g}), \quad a_2 = \frac{l\kappa_b^{(0)}}{l_b\kappa_s^{(0)}},$$

$$F_1 = \frac{AI_0}{2\beta\kappa_s}(1 - \beta l - E), \quad F_2 = \frac{AI_0}{2\beta\kappa_s}(1 - (1 + \beta l)E).$$

Substituting $\delta_{2g} = \delta_{2s} = \delta_{2b} = 0$ into Eqs. (15) and (16), we obtain the results predicted by the linear

theory [13]:

$$\Theta_L = \frac{I_0}{2\kappa_s^{(0)}\beta} \frac{(1 - E)\left(l\beta\frac{a_2}{a_1} - 1 \right) + \beta l}{1 + a_1 + a_2}, \tag{17}$$

$$W_L = \frac{I_0 l_b}{2\kappa_b^{(0)}}(1 - E) - a_2\Theta_L.$$

It should be emphasized that, in the case under study, it is important to know the temperature at the sample–gas interface, since it is the temperature variations at this interface that produce the PA signal detected by a highly sensitive microphone. The values of Θ_0 and W_0 were numerically calculated for ebonite ($l = 0.001$ m) in contact with air ($l_g = 0.005$ m). A stainless-steel plate ($l_b = 0.002$ m) was used as a substrate. This cell geometry is typical of PA cells used in studying condensed media. The processing of the experimental temperature dependences of the thermal conductivities of air [17], ebonite [18], and stainless steel [17] gives the following parameter values: $\delta_{2g} = 2.39 \times 10^{-3}$ K^{-1}, $\delta_{2s} = 6.41 \times 10^{-3}$ K^{-1}, and $\delta_{2b} = 0.24 \times 10^{-3}$ K^{-1}. Here, the thermal conductivities of these media at $T = 300$ K are taken to be $\kappa_g^{(0)} = 0.025668$ W/(m K), $\kappa_b^{(0)} = 14.7796$ W/(m K), and $\kappa_s^{(0)} = 0.1266$ W/(m K), respectively. The values $\beta = 10^7$ m^{-1} and $R = 0.2$ are taken from [19].

The dependence of Θ_0 on I_0 obtained by numerically solving Eqs. (15) and (16) is shown in Fig. 2 (curve *2*). For the sake of comparison, Θ_L versus I_0 is also plotted (curve *1*). It can be seen that (i) the linear dependence of Θ_0 on I_0 holds for small values of I_0 (which corresponds to the absence of thermal nonlinearity and, hence, complies with the results of [13]) and (ii) the dependence of Θ_0 on I_0 substantially deviates from linear at $I_0 \geq 1$ W/cm^2 and then passes to a power law dependence with Θ_0 being lower than Θ_L. The latter fact is due to the increase in the thermal conductivity with increasing temperature and a resulting increase in the efficiency of heat transfer to the surrounding medium.

Calculations show that the inequality $W_0/\Theta_0 < 0.1$ holds. Then, the terms with W_0 in Eq. (15) can be omitted and the solution to the resulting quadratic equation is

$$\Theta_0 = \frac{\left[\sqrt{4b_{11}|F_1| + b_{12}^2} - b_{12} \right]}{2b_{11}}.$$

Since $b_{12} \approx 1$ in this case, then, for $4b_{11}|F_1| \gg 1$, we obtain a square-root dependence of Θ_0 on I_0:

$$\Theta_0 \approx \sqrt{\frac{|F_{11}|}{b_{11}}} \approx \sqrt{\frac{AI_0 l}{2\kappa_s^{(0)} b_{11}}}$$

(here, we also take into account that $\beta l \gg 1$), which is confirmed by the shape of curve *2* in Fig. 2.

3. NONSTEADY TEMPERATURE FIELD: A LINEAR MODEL

When a sample in a PA cell is irradiated with a low-power source (e.g., a helium–neon laser), the sample is heated insignificantly, so that the thermophysical parameters of the medium remain unchanged. In this case, nonlinear differential heat conduction equations (3)–(5) transform into the following linear equations:

$$\frac{\partial T_g}{\partial t} - \chi_g \frac{\partial^2 T_g}{\partial x^2} = 0, \quad 0 \le x \le l_g, \tag{18}$$

$$\frac{\partial T_s}{\partial t} - \chi_s \frac{\partial^2 T_s}{\partial x^2} = \frac{A\beta I_0}{2\rho C_p} \exp(\beta x), \quad -l \le x \le 0, \tag{19}$$

$$\frac{\partial T_b}{\partial t} - \chi_b \frac{\partial^2 T_b}{\partial x^2} = 0, \quad -(l + l_b) \le x \le -l. \tag{20}$$

Here, the superscript (0) and primes are omitted for brevity.

The conditions of heat flux continuity at the boundaries also become linear:

$$\kappa_g \frac{dT_g}{dx}\bigg|_{x=0} = \kappa_s \frac{\partial T_s}{\partial x}\bigg|_{x=0}, \tag{21}$$

$$\kappa_b \frac{dT_b}{dx}\bigg|_{x=-l} = \kappa_s \frac{dT_s}{dx}\bigg|_{x=-l}.$$

The set of Eqs. (18)–(20) together with boundary conditions (6) and (21) and the initial conditions

$$T_g(0, x) = T_s(0, x) = T_b(0, x) = 0 \tag{22}$$

present a mathematical formulation of the time-dependent problem of determining a TF in a one-dimensional PA cell in terms of a linear model. This problem can be solved by applying integral Laplace transformation with respect to time:

$$\tilde{T}_i(p, x) = \int_0^\infty \exp(-pt) T(t, x) dt. \tag{23}$$

Substituting Eq. (23) into Eqs. (18)–(20) and taking into account initial conditions (22), we have

$$\frac{d^2\tilde{T}_g}{dx^2} - q_g^2 \tilde{T}_g = 0, \quad 0 \le x \le l_g, \tag{24}$$

$$\frac{d^2\tilde{T}_g}{dx^2} - q_s^2 \tilde{T}_s = -\frac{A\beta I_0}{2p\kappa_s} \exp(\beta x), \quad -l \le x \le 0, \tag{25}$$

$$\frac{d^2\tilde{T}_b}{dx^2} - q_b^2 \tilde{T}_b = 0, \quad -(l + l_b) \le x \le -l, \tag{26}$$

where $q_i^2 = p/\chi_i$ and $i = g, s, b$.

The solutions to Eqs. (24) and (26) are given by the expressions

$$\tilde{T}_g(p, x) = \tilde{T}_s(p, 0) \frac{\sinh\lfloor q_g(l_g - x)\rfloor}{\sinh(l_g q_g)},$$

$$\tilde{T}_b(p, x) = \tilde{T}_s(p, -l) \frac{\sinh[q_b(x + l + l_b)]}{\sinh(q_b l_b)}, \tag{27}$$

which satisfy boundary conditions (6). The solution to Eq. (25) can be found by the method of variation of constants:

$$\tilde{T}_s(p, x) = \gamma_1 \exp(q_s x) + \gamma_2 \exp(-q_s x) - \frac{A\beta I_0 \exp(\beta x)}{2\kappa_s(\beta^2 - q_s^2)p}. \tag{28}$$

Substituting γ_1 and γ_2 derived from Eqs. (21) into expression (28), we obtain

$$\tilde{T}_s(p, x) = \left[\frac{\Delta_1(p)}{q_s\Delta(p)} \exp(q_s x) - \frac{\Delta_2(p)}{q_s\Delta(p)} \exp(-q_s x) + \exp(\beta x) \right] \frac{A\beta I_0}{2\kappa_s p(q_s^2 - \beta^2)}. \tag{29}$$

Here, the following notation is introduced:

$$\Delta(p) = (\beta_g\beta_b + 1)(\exp(-q_s l) - \exp(q_s l)) - (\beta_g + \beta_b)(\exp(q_s l) + \exp(-q_s l));$$

$$\Delta_1(p) = (q_s\beta_g + \beta\beta_b)\exp(q_s l) + (\beta + q_s\beta_g\beta_b)(\exp(q_s l) - \exp(-\beta l)) + (\beta\beta_g + q_s\beta_b)\exp(-\beta l);$$

$$\Delta_2(p) = (\beta - q_s\beta_g\beta_b)(\exp(-\beta l) - \exp(-q_s l)) + (\beta\beta_g - q_s\beta_b)\exp(-\beta l) - (q_s\beta_g - \beta\beta_b)\exp(-q_s l);$$

$$\beta_i = c_i \coth(l_i q_i); \quad c_i = \frac{\kappa_i}{\kappa_s}\sqrt{\frac{\chi_s}{\chi_i}}; \quad i = g, b.$$

To determine $T(t, x)$, it is necessary to perform an inverse Laplace transformation of Eq. (29), i.e., to calculate the integral

$$T_s(t, x) = \frac{A\beta I_0}{2\kappa_s 2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{\exp(pt)}{p(q_s^2 - \beta^2)} dp \left[\frac{1}{q_s} \frac{\Delta_1(p)}{\Delta(p)} \right.$$

$$\left. \times \exp(q_s x) - \frac{\Delta_2(p)}{q_s\Delta(p)} \exp(-q_s x) + \exp(\beta x) \right]. \tag{30}$$

To do this, we employ the convolution theorem, for which we represent the first two terms in the integrand in the form

$$\frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} f_{1(i)}(p) f_{2(i)}(p) \exp(pt) dp, \tag{31}$$

where

$$f_{1(1)}(p) = \frac{\exp(-q_s|x|)}{q_s}, \quad f_{1(2)}(p) = \frac{\exp[-(l-|x|)q_s]}{q_s},$$

$$f_{2(1)}(p) = \frac{\Delta_1(p)}{\Delta(p)(q_s^2-\beta^2)p}, \quad f_{2(2)}(p) = \frac{\Delta_2(p)\exp(q_sl)}{\Delta(p)(q_s^2-\beta^2)p}.$$

The functions (see, e.g., [20])

$$f_{1(1)}(\tau) = \left(\frac{\chi}{\pi\tau}\right)^{1/2}\exp\left(-\frac{|x|^2}{4\chi\tau}\right),$$

$$f_{1(2)}(\tau) = \sqrt{\frac{\chi}{\pi\tau}}\exp\left(-\frac{(l-|x|)^2}{4\chi\tau}\right),$$

(32)

$$f_{2(1)}(t) = \frac{1}{\beta^2}[\exp(\chi\beta^2 t)\psi(p_1) - \psi_1(0)],$$

$$f_{2(2)}(t) = \frac{1}{\beta^2}[\exp(\chi\beta^2 t)\phi(p_1) - \phi_1(0)]$$

(33)

are the result of integrating with respect to $f_{1(i)}(p)$ and $f_{2(i)}(p)$, where

$$p_1 = \chi\beta^2, \quad \psi(p_1) = \lim_{p\to p_1}\frac{\Delta_1(p)}{\Delta(p)},$$

$$\phi(p_1) = \lim_{p\to p_1}\frac{\Delta_2(p)\exp(q_sl)}{\Delta(p)},$$

$$\psi_1(0) = \lim_{p\to 0}\frac{\Delta_1(p)}{\Delta(p)}, \quad \phi_1(0) = \lim_{p\to 0}\frac{\Delta_2(p)\exp(q_sl)}{\Delta(p)}.$$

The functions $\psi(p_1)$ and $\phi(p_1)$ are easy to calculate. However, $\psi_1(0)$ and $\phi_1(0)$ should be calculated with caution; namely, it is necessary to expand both the numerators and denominators in powers of q_s and keep the first nonvanishing terms. As a result, we obtain

$$\psi(p_1) = \frac{\beta B_1}{B}, \quad \psi_1(0) = -\frac{B_{01}}{2B_{00}},$$

$$\phi(p_1) = 0, \quad \phi_1(0) = -\frac{B_{02}}{2B_{00}},$$

$$B = (1 + \beta_{og}\beta_{ob})\sinh(-\beta l) - (\beta_{og}+\beta_{ob})\cosh(\beta l),$$

$$B_{00} = c_g c_b l + c_g l_b\sqrt{\frac{\chi_s}{\chi_b}} + c_b l_b\sqrt{\frac{\chi_s}{\chi_g}}, \quad (34)$$

$$B_1 = (1 + \beta_{og}B_{ob})\cosh(\beta l) + (\beta_{og}+\beta_{ob})\sinh(\beta l), (35)$$

$$B_{01} = c_g c_b(1 + \exp(-\beta l))$$

$$+ c_b l_g\beta\sqrt{\frac{\chi_s}{\chi_g}} + c_g\sqrt{\frac{\chi_s}{\chi_b}}\beta l_b\exp(-\beta l), \quad (36)$$

$$B_{02} = c_g c_b(1 - \exp(-\beta l))$$

$$+ \beta l_g c_b\sqrt{\frac{\chi_s}{\chi_g}} + c_g\beta l_b\sqrt{\frac{\chi_s}{\chi_b}}\exp(-\beta l),$$

$$\beta_{og} = c_g\coth\left(\beta l_g\sqrt{\frac{\chi_s}{\chi_g}}\right),$$

$$\beta_{0b} = c_b\coth\left(\beta l_b\sqrt{\frac{\chi_s}{\chi_b}}\right).$$

(37)

Taking into consideration Eqs. (32) and (33), we obtain the final expression for TF:

$$T_s(t,x) = \frac{AI_0}{2\beta\kappa_s}\Big[\exp(\beta x)(\exp(\chi_s\beta^2 t) - 1)$$

$$+ \exp(\chi_s\beta^2 t)\frac{\beta B_1}{B}I_1 + \frac{B_{01}I_2}{2B_{00}} - \frac{B_{02}I_3}{2B_{00}}\Big].$$

(38)

Here, the following notation is used:

$$I_1 = \left(\frac{\chi_s}{\pi}\right)^{1/2}\int_o^t \frac{\exp\left(-\frac{|x|^2}{4\chi_s\tau}\right) - \chi_s\beta^2\tau}{\sqrt{\tau}}d\tau, \quad (39)$$

$$I_2 = \sqrt{\frac{\chi_s}{\pi}}\int_o^t \frac{\exp\left(-\frac{|x|^2}{4\chi_s\tau}\right)}{\sqrt{\tau}}d\tau,$$

(40)

$$I_3 = \sqrt{\frac{\chi_s}{\pi}}\int_o^t \frac{\exp\left(-\frac{(l-|x|)^2}{4\chi_s\tau}\right)}{\sqrt{\tau}}d\tau.$$

Figures 3 and 4 show the results of numerical calculations for ebonite at two values of I_0 (curves *1* and *2*). The geometric parameters and the thermal conductivities at $T = 300$ K are specified in Section 2. The other parameter values at this temperature are taken to be $\rho_g = 1.29$ kg/m^3, $\rho_b = 7700$ kg/m^3, $\rho_s = 1200$ kg/m^3, $C_{pg} = 1008$ J/(kg K), $C_{pb} = 460$ J/(kg K), and $C_{ps} = 1319.2$ J/(kg K) [17]. As was expected, the TF is seen to relax to an equilibrium state. However, these results hold true only in the linear approximation, in which the thermophysical parameters of the sample, gas, and substrate materials are assumed to be constant.

4. NONSTEADY TEMPERATURE FIELD: A NONLINEAR MODEL

Highly absorptive media placed into a PA cell and irradiated by a moderate-intensity laser beam (~1 W/cm^2) gain a considerable amount of heat. In this case, the temperature of the medium increases substantially and the effect of the temperature dependence of the thermophysical parameters of the media begins to reveal itself in the PA signal. The nonsteady TF can be determined from the set of nonlinear heat conduction
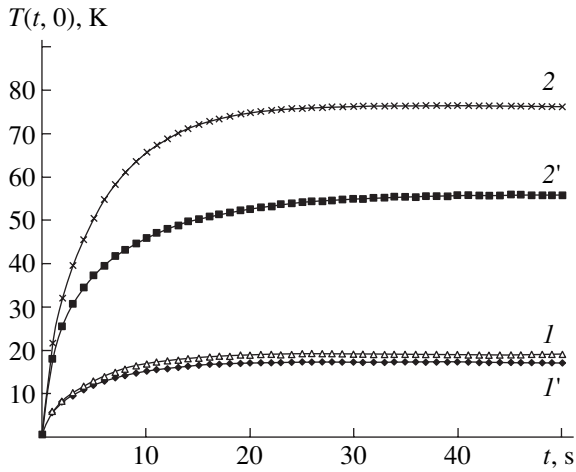
Fig. 3. Evolution of the temperature at the ebonite–air boundary ($x = 0$) in a PA cell at $I_0 = $ (*1*, *1'*) 0.5 and (*2*, *2'*) 2 W/cm² in terms of (*1*, *2*) linear and (*1'*, *2'*) nonlinear models.
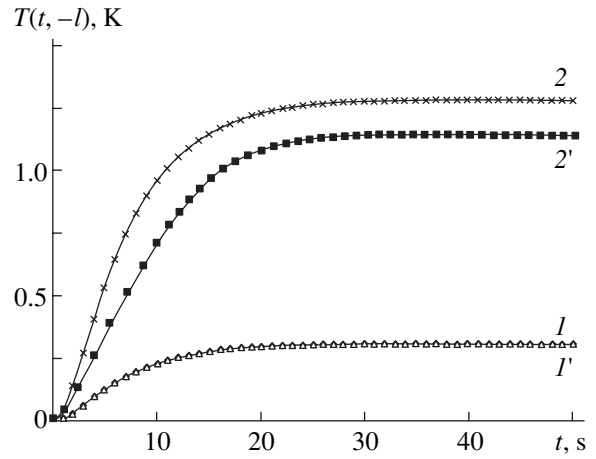

Fig. 4. Evolution of the temperature at the substrate–sample interface ($x = -l$) in a PA cell. The parameters are the same as in Fig. 3.

equations (3)–(5) for all the layers in the PA cell. This system should take into account the temperature dependences of the densities, specific heats, and thermal conductivities of the layers. These dependences are evidently determined by the physical properties of a given system, and each particular case requires a separate analysis.

As before, we will consider a three-layer (air–ebonite–stainless steel) system with the temperature dependences of the thermal conductivities specified in Section 2. Since the coefficient of thermal expansion of solids is as low as $\approx 10^{-6}$ K^{-1}, one may ignore the temperature-induced decrease in the density of the sample and substrate; i.e., it can be assumed that $\rho_s(T) = \rho_s(T_0)$ and $\rho_b(T) = \rho_b(T_0)$. The gas density also remains constant in view of the fixed volume of the gas-filled part of the cell: $\rho_g = \rho(T_0)$. The temperature factor of the air specific heat is $\delta_{1g} = 0.19 \times 10^{-3}$ K^{-1} at $C_{pg}(300 \text{ K}) \approx 1008$ J/(kg K) [17]. According to the dependences of W on t calculated in the previous section for different values of I_0 (Fig. 4, curves *1*, *2*), the substrate is heated insignificantly, so that the quantity C_{pb} can be considered constant. In ebonite, the dependence of C_{ps} on T was estimated from the curves $\kappa(T)$ and $\chi(T)$ presented in [18] and appeared to be well approximated by $C_{ps}(300 \text{ K}) \approx 1319$ J/(kg K) and $\delta_{1s} = 15.7 \times 10^{-3}$ K^{-1} within the 300- to 360-K temperature range. With allowance for the above factors, substituting Eq. (1) into Eqs. (3) and (4) results in the following set of equations:

$$(1 + \delta_{1g}T_g)\frac{\partial T_g}{\partial t} = \chi_g^{(0)}\frac{\partial}{\partial x}\left[(1 + \delta_{2g}T_g)\frac{\partial T_g}{\partial x}\right], \quad (41)$$
$$0 \leq x \leq l_g,$$

$$(1 + \delta_{1s}T_s)\frac{\partial T_s}{\partial t} = \chi_s^{(0)}\frac{\partial}{\partial x}\left[(1 + \delta_{2s}T_s)\frac{\partial T_s}{\partial x}\right]$$
$$+ \frac{A\beta I_0 \exp(\beta x)}{2\rho_s C_p}, \quad -l \leq x \leq 0, \quad (42)$$

$$\frac{\partial T_b}{\partial t} = \chi_b^{(0)}\frac{\partial^2 T_b}{\partial x^2}, \quad -(l + l_b) \leq x \leq -l, \quad (43)$$

which can be solved numerically by the finite difference method. The time dependences of the temperature of the irradiated ebonite surface $\Theta = T_s(0, t)$ and the sample–substrate interface $W = T_s(-l, t)$ are shown in Figs. 3 and 4, respectively, for two values of I_0 (curves *1'* and *2'*). It can be seen that the temperature dependence of the thermophysical parameters results in a decrease in the equilibrium temperature (which complies with the conclusions drawn in Section 3) and somewhat faster relaxation to a steady TF. A comparison between the two curves shows that $\Theta/W \approx 30$–50 within the entire time interval under study; i.e., the heating of the substrate–sample interface is much lower than that of the irradiated sample surface. This estimate is important for the development of a nonlinear theory of the PA effect since it allows one to restrict oneself to the linear approximation in Eq. (5).

CONCLUSIONS

Exact expressions for a nonsteady TF in a one-dimensional three-layer PA in a linear approximation are obtained. The corresponding nonlinear problem implying that the thermophysical parameters of all the layers in the cell are temperature-dependent is solved numerically. An exact solution to the nonlinear tome-independent problem is obtained. It is found that, due to

thermal nonlinearity, the initially linear dependence of the temperature of the irradiated sample surface on the radiation intensity transforms into a power-law dependence as the intensity increases. It is shown that the heating of the substrate surface contacting a highly absorptive low-conductivity sample is at least one order of magnitude lower than the heating of the irradiated sample surface. This fact is of importance since it allows one to ignore the effect of a substrate in the nonlinear theory of the PA effect for highly absorptive media in the above geometry.

ACKNOWLEDGMENTS

REFERENCES

1. S. V. Egerev, L. M. Lyamshev, and O. V. Puchenkov, Usp. Fiz. Nauk **160** (9), 111 (1990) [Sov. Phys. Usp. **33**, 739 (1990)].

2. A. Mandelis, Phys. Today **53** (8), 29 (2000).

3. N. A. Pankratov, Opt. Zh. **61** (5), 3 (1994).

4. U. Madvaliev, Zh. Tekh. Fiz. **50**, 880 (1980) [Sov. Phys. Tech. Phys. **25**, 530 (1980)].

5. K. L. Muratikov and A. L. Glazov, Zh. Tekh. Fiz. **71** (6), 110 (2001) [Tech. Phys. **46**, 749 (2001)].

6. A. Yu. Luk'yanov and A. A. Pogorelko, Zh. Tekh. Fiz. **72** (5), 72 (2002) [Tech. Phys. **47**, 585 (2002)].

7. K. L. Muratikov, A. L. Glazov, D. N. Rouz, and D. E. Dumar, Pis'ma Zh. Tekh. Fiz. **28** (9), 48 (2002) [Tech. Phys. Lett. **28**, 377 (2002)].

8. V. É. Gusev and A. A. Karabutov, *Laser Optoacoustics* (Nauka, Moscow, 1991).

9. Y. N. Rajakarunanayake and H. K. Wickramasinghe, Appl. Phys. Lett. **48**, 218 (1986).

10. C. Wang and P. Li, J. Appl. Phys. **74** (9), 5713 (1993).

11. A. Mandelis, A. Salnik, J. Opsal, *et al.*, J. Appl. Phys. **85**, 1811 (1999).

12. V. Gusev, A. Mandelis, and R. Bleiss, Mater. Sci. Eng., B **26**, 111 (1994).

13. A. Rosencwaig and A. Gersho, J. Appl. Phys. **47**, 64 (1976).

14. F. A. McDonald and G. C. Wetsel, Jr., J. Appl. Phys. **49**, 2313 (1978).

15. F. A. McDonald and G. C. Wetsel, Jr., Phys. Acoust. **18**, 168 (1988).

16. Y. Gurevich, G. Logvinov, Luis N. de Rivera, *et al.*, Rev. Sci. Instrum. **74**, 441 (2003).

17. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991; CRC, Boca Raton, 1997).

18. L. L. Vasil'ev and S. A. Tanaeva, *Thermal Physical Properties of Porous Materials* (Nauka i Tekhnika, Minsk, 1971).

19. N. G. Basov, V. A. Boĭko, O. N. Krokhin, *et al.*, Zh. Tekh. Fiz. **38**, 1973 (1968) [Sov. Phys. Tech. Phys. **13**, 1581 (1968)].

20. H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon, Oxford, 1959; Nauka, Moscow, 1964).

Translated by A. Sidorova

<div align="center">

THEORETICAL
AND MATHEMATICAL PHYSICS

Simulation of Wave Scattering by Bodies with an Absorbing Coating and Black Bodies

</div>

A. G. Kyurkchan and D. B. Demin

Moscow Technical University of Communication and Informatics, Moscow, 111024 Russia
e-mail: kyurkchan@mtuci2.ru
Received June 30, 2003

Abstract—A new effective method for solving three-dimensional problems of electromagnetic wave diffraction by impedance bodies with irregular boundaries is proposed. The method offers a high rate of convergence. Examples of solving the problems of wave scattering by bodies of revolution are given, and results illustrating the rate of convergence of a computational algorithm for bodies of various shapes are presented. The impedance approximation is shown to be valid for simulation of scattering characteristics of bodies with an insulating coating even when the boundary has irregularities and the refractive index of the coating is not too high. Various ways of characterizing "black bodies" and the results of studying their scattering characteristics are discussed. © 2004 MAIK "Nauka/Interperiodica".

INTRODUCTION

The problem of simulating the scattering characteristics of nonreflecting or "black" bodies is a classical problem of electrodynamics and diffraction theory. A fundamental difficulty of this problem is that a so-called black body cannot be described in terms of the rigorous boundary-value problem of the diffraction theory [1]. Therefore, the scattering characteristics of such bodies are described with various approximations [1−3]. The associated methods of description (models) can be conventionally divided into two classes. The first one covers so-called structural models, where the production of the absorbing coating is discussed; in the second one (speculative models), the realization of any particular model is not the case in point. The Macdonald concept [2] belongs to the second class. According to this concept, the black body surface is an ideal electrical conductor ($\varepsilon_r = -i\infty$ and $\mu_r = 1$) and an ideal magnetic ($\varepsilon_r = 1$ and $\mu_r = -i\infty$) conductor simultaneously. The field scattered by such a body is calculated as the half-sum of the fields scattered by ideal electrical and magnetic conductors of the same shape. The well-known Sommerfeld model [1] may be classified as a structural model. This model implies that the low-reflection coating must have a relative permittivity ε_r equal to the relative permeability μ_r. In addition, the coating must absorb electromagnetic radiation; i.e., ε_r and μ_r must be complex and have large imaginary parts.

Until recently, most of the results on wave scattering by black bodies have been obtained with approximate (asymptotic) approaches, such as the method of physical optics and its generalizations [3]. These are inapplicable when the dimensions of the body are comparable to the wavelength. In this case, the rigorous approaches

(adopted for black body description) are needed. The results of such investigations are few [2, 4].

In this paper, we develop an effective method for solving the problem of electromagnetic wave diffraction by bodies with an impedance boundary (the pattern equation method, PEM). The method makes it possible to study the characteristics of wave scattering by bodies with a magnetodielectric coating. The results of this research are reported.

PROBLEM DEFINITION AND FORMULATION OF THE PATTERN EQUATION METHOD

Consider the 3D problem of scattering a primary monochromatic electromagnetic field $\mathbf{E}^0$, $\mathbf{H}^0$ by a finite target bounded by a surface S. Let the following boundary condition be set on the surface S:

$$(\mathbf{n} \times \mathbf{E})|_S = -Z[\mathbf{n} \times (\mathbf{n} \times \mathbf{H})]|_S. \qquad (1)$$

Here, Z is the surface impedance [5]; $\mathbf{n}$ is the unit normal to the surface S; $\mathbf{E} = \mathbf{E}^0 + \mathbf{E}^1$ and $\mathbf{H} = \mathbf{H}^0 + \mathbf{H}^1$ specify the total field; and $\mathbf{E}^1$ and $\mathbf{H}^1$ specify the secondary (diffraction) field that satisfies the homogeneous Maxwell equations

$$\nabla \times \mathbf{E}^1 = -ik\zeta \mathbf{H}^1, \quad \nabla \times \mathbf{H}^1 = \frac{ik}{\zeta}\mathbf{E}^1,$$

everywhere outside S and also a condition at infinity, e.g., of the form

$$\left(\mathbf{E}^1 \times \frac{\mathbf{r}}{r}\right) + \zeta\mathbf{H} = o\left(\frac{1}{r}\right), \quad \left(\mathbf{H}^1 \times \frac{\mathbf{r}}{r}\right) - \frac{1}{\zeta}\mathbf{E} = o\left(\frac{1}{r}\right),$$

$$r \equiv |\mathbf{r}| \longrightarrow \infty.$$

Here, $k = \omega\sqrt{\varepsilon\mu}$ is the wavenumber and $\zeta = \sqrt{\mu/\varepsilon}$ is the impedance of the medium. According to the PEM standard scheme [6, 7], we will search for the scattering pattern (wave field diagram), i.e., a function that relates the diffraction field to the angles θ and φ of the spherical coordinate system (r, θ, φ) in the far zone ($kr \gg 1$). In this zone, the asymptotic relationships of the form

$$\mathbf{E}^1 = \frac{\exp(-ikr)}{r}\mathbf{F}^E(\theta, \varphi) + O\left(\frac{1}{(kr)^2}\right),$$

$$\mathbf{H}^1 = \frac{\exp(-ikr)}{r}\mathbf{F}^H(\theta, \varphi) + O\left(\frac{1}{(kr)^2}\right),$$

where $\mathbf{F}^E$ and $\mathbf{F}^H$ are the patterns of the electrical and magnetic fields, respectively, are satisfied.

Let us consider briefly the derivation of an integro-operator equation of the PEM. For simplicity, only the case $Z = 0$ will be discussed. The starting point in this consideration is the Sommerfeld–Weyl representation of a wave field (e.g., magnetic field) in the form of the generalized integral of plane waves [8]:

$$\mathbf{H}^1 = \frac{1}{2\pi i}\int_0^{2\pi} d\beta \int_0^{\pi/2+i\infty} \exp(-ikr\cos\alpha) \tag{2}$$

$$\times \hat{\mathbf{F}}^H(\theta, \varphi; \alpha, \beta)\sin\alpha\, d\alpha,$$

where

$$\hat{\mathbf{F}}^H(\theta, \varphi; \alpha, \beta)$$

$$= \frac{ik^2}{4\pi}\int_S ((\mathbf{n'}\times\mathbf{H})\times\hat{\mathbf{p}})\exp(ik\hat{\mathbf{p}}\cdot\mathbf{r'})ds' \tag{3}$$

is the scattering pattern in the coordinate system rotated so that the $0Z$ axis is directed to the point of observation.

In (3), $\mathbf{p} = \{\sin\alpha\cos\beta, \sin\alpha\sin\beta, \cos\alpha\}$ and $\hat{\mathbf{p}} = A^T\mathbf{p}$, where

$$A = \begin{pmatrix} -\sin\varphi & \cos\varphi & 0 \\ -\cos\varphi\cos\theta & -\sin\varphi\cos\theta & \sin\theta \\ \cos\varphi\sin\theta & \sin\varphi\sin\theta & \cos\theta \end{pmatrix}$$

is the matrix of rotation of the coordinate system.

Integral (2) converges absolutely and uniformly in $R^3/\bar{B}$, where $\bar{B}$ is the convex hull of scattered field singularities [9]. Taking into account that $\mathbf{H} = \mathbf{H}^0 + \mathbf{H}^1$, we find the integral operator equation desired by substituting (2) into (3):

$$\mathbf{F}^H(\xi, \eta) = \mathbf{F}_0^H(\xi, \eta) + \frac{k^2}{4\pi i}\int_S \mathbf{q}\times\mathbf{n}\frac{1}{2\pi i}$$

$$\times \int_0^{2\pi}\int_0^{\pi/2+i\infty} \exp(-ik\rho\cos\alpha)\hat{\mathbf{F}}^H(\theta, \varphi; \alpha, \beta) \tag{4}$$

$$\times \exp(ik\mathbf{q}\boldsymbol{\rho}(\theta, \varphi))\sin\alpha\, d\alpha\, d\beta\, ds,$$

where

$$\mathbf{q} = \{\sin\xi\cos\eta, \sin\xi\sin\eta, \cos\xi\},$$

$$\mathbf{F}_0^H(\xi, \eta) = \frac{ik^2}{4\pi}\int_S [(\mathbf{n}\times\mathbf{H}^0)\times\mathbf{q}]\exp(ik\mathbf{q}\cdot\boldsymbol{\rho'})ds'.$$

Expanding the pattern in some basis, substituting the expansion into Eq. (4), and projecting the right- and left-hand sides of the equation onto some (generally, another) basis on the unit sphere, we reduce the problem to an infinite algebraic system. This method is mathematically justified if $\bar{B}$ is located inside a scatterer [6, 7, 10]. However, in the vector form with impedance boundary conditions, such a computational scheme is cumbersome. Therefore, we will consider another method (which is mathematically simpler while more difficult to substantiate) of deriving a PEM algebraic set of equations for the coefficients of expansion of the pattern in spherical harmonics [11, 12]. With this approach, the integro-operator equation is unnecessary.

REDUCTION OF THE BOUNDARY-VALUE PROBLEM TO A SET OF ALGEBRAIC EQUATIONS

It is known (see, e.g., [12]) that

$$\mathbf{F}^E(\theta, \varphi) = -\sum_{n=1}^{\infty}\sum_{m=-n}^{n} a_{nm}i^n(\mathbf{i}_r\times\boldsymbol{\Phi}_n^m(\theta, \varphi))$$

$$-\sum_{n=1}^{\infty}\sum_{m=-n}^{n} b_{nm}i^n\zeta\boldsymbol{\Phi}_n^m(\theta, \varphi), \tag{5}$$

$$\mathbf{F}^H(\theta, \varphi) = -\sum_{n=1}^{\infty}\sum_{m=-n}^{n} a_{nm}i^n\frac{1}{\zeta}\boldsymbol{\Phi}_n^m(\theta, \varphi)$$

$$-\sum_{n=1}^{\infty}\sum_{m=-n}^{n} b_{nm}i^n(\mathbf{i}_r\times\boldsymbol{\Phi}_n^m(\theta, \varphi)), \tag{6}$$

where

$$\boldsymbol{\Phi}_n^m(\theta, \varphi) = \mathbf{r}\times\nabla P_n^m(\cos\theta)\exp(im\varphi). \tag{7}$$

Thus, our aim is to find an algebraic system for the coefficients a_{nm} and b_{nm}.

The wave fields $\mathbf{E}^1$ and $\mathbf{H}^1$ can be represented as the expansions

$$\mathbf{E}^1 = \sum_{n=1}^{\infty}\sum_{m=-n}^{n} \{a_{nm}[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{r}\psi_n^m)] \tag{8}$$
$$-ik\zeta b_{nm}[\boldsymbol{\nabla}\times(\mathbf{r}\psi_n^m)]\},$$

$$\mathbf{H}^1 = \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left\{\frac{ik}{\zeta}a_{nm}[\boldsymbol{\nabla}\times(\mathbf{r}\psi_n^m)] \right.$$
$$\left. + b_{nm}[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{r}\psi_n^m)]\right\}, \tag{9}$$

$$\psi_n^m = h_n^{(2)}(kr)P_n^m(\cos\theta)\exp(im\varphi), \tag{10}$$

where $h_n^{(2)}$ are the Hankel spherical functions of the second kind and P_n^m are the associated Legendre polynomials.

The starting point for further analysis is the representation of the coefficients a_{nm} and b_{nm} via the boundary values of the vector field. The expressions for these coefficients, by analogy with [11], can be found from the relationships

$$\mathbf{E}^1 = \int_S\left\{\frac{\zeta}{ik}[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{I}^e G_0)] - Z[\boldsymbol{\nabla}\times(\mathbf{I}^m G_0)]\right\}ds', \tag{11}$$

$$\mathbf{H}^1 = \int_S\left\{\frac{Z}{ik\zeta}[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{I}^m G_0)] + [\boldsymbol{\nabla}\times(\mathbf{I}^e G_0)]\right\}ds', \tag{12}$$

where

$$\mathbf{I}^e = (\mathbf{n}\times\mathbf{H})|_S, \quad \mathbf{I}^m = (\mathbf{n}H_n - \mathbf{H})|_S,$$
$$G_0 = G_0(\mathbf{r},\mathbf{r}') = \frac{\exp(-ik|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{13}$$

is the Green's function for free space (the fundamental solution of the Helmholtz scalar equation).

Then, using (8)–(13), one finds

$$a_{nm} = \frac{k}{4\pi i}N_{nm}\int_S\left\{\frac{\zeta}{ik}\mathbf{I}^e[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{r}\bar\chi_n^m)] \right.$$
$$\left. - Z\mathbf{I}^m[\boldsymbol{\nabla}\times(\mathbf{r}\bar\chi_n^m)]\right\}ds, \tag{14}$$

$$b_{nm} = \frac{k}{4\pi i}N_{nm}\int_S\left\{\mathbf{I}^e[\boldsymbol{\nabla}\times(\mathbf{r}\bar\chi_n^m)] \right.$$
$$\left. + \frac{Z}{ik\zeta}\mathbf{I}^m[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{r}\bar\chi_n^m)]\right\}ds, \tag{15}$$

where

$$\bar\chi_n^m(r) = j_n(kr)P_n^m(\cos\theta)\exp(-im\varphi),$$
$$N_{nm} = \frac{2n+1(n-m)!}{n(n+1)(n+m)!},$$

j_n are the spherical Bessel functions, and the bar means complex conjugation.

Relationships (14) and (15) are the starting point for finding the algebraic system desired. Let us introduce the notation

$$\mathbf{H}_{nm}^h = \boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{r}\psi_n^m), \quad \mathbf{H}_{nm}^e = \frac{ik}{\zeta}\boldsymbol{\nabla}\times(\mathbf{r}\psi_n^m). \tag{16}$$

From (9) and (16), we have

$$\mathbf{H} = \mathbf{H}_0 + \sum_{q=1}^{\infty}\sum_{p=-q}^{q}(a_{qp}\mathbf{H}_{qp}^e + b_{qp}\mathbf{H}_{qp}^h). \tag{17}$$

Then, in view of (14), (15), and (17), we find the PEM system:

$$\begin{cases}
a_{nm} = a_{nm}^0 + \sum_{q=1}^{\infty}\sum_{p=-q}^{q}(G_{nm,qp}^{11}a_{qp} + G_{nm,qp}^{12}b_{qp}) \\
b_{nm} = b_{nm}^0 + \sum_{q=1}^{\infty}\sum_{p=-q}^{q}(G_{nm,qp}^{21}a_{qp} + G_{nm,qp}^{22}b_{qp}) \\
n = 1, 2, \ldots; \quad |m| \le n.
\end{cases} \tag{18}$$

Here,

$$a_{nm}^0 = a_{nm}^{00} + Za_{nm}^{z0}; \quad b_{nm}^0 = b_{nm}^{00} + Zb_{nm}^{z0};$$
$$G_{nm,qp}^{ij} = G_{nm,qp}^{0ij} + ZG_{nm,qp}^{zij}, \tag{19}$$

where the second superscripts 0 and z denote the corresponding values at $Z = 0$ and additional terms that result from the fact that the impedance is nonzero. The components of (19) are given by

$$a_{nm}^{00} = \frac{-\zeta}{4\pi}N_{nm}\int_S(\mathbf{n}\times\mathbf{H}^0)\overline{\mathbf{e}_{nm}^e}ds,$$

$$b_{nm}^{00} = \frac{\zeta}{4\pi}N_{nm}\int_S(\mathbf{n}\times\mathbf{H}^0)\overline{\mathbf{h}_{nm}^e}ds,$$

$$a_{nm}^{z0} = \frac{-\zeta}{4\pi}N_{nm}\int_S[\mathbf{n}\times(\mathbf{n}\times\mathbf{H}^0)]_S\overline{\mathbf{h}_{nm}^e}ds, \tag{20}$$

$$b_{nm}^{z0} = -\frac{1}{4\pi\zeta} N_{nm} \int_S [\mathbf{n} \times (\mathbf{n} \times \mathbf{H}^0)]_S \overline{\mathbf{e}_{nm}^e} ds.$$

The matrix elements of the PEM system at $Z = 0$ have the form [11]

$$G_{nm,qp}^{011} = -\frac{\zeta}{4\pi} N_{nm} \int_S (\mathbf{n} \times \mathbf{H}_{qp}^e) \overline{\mathbf{e}_{nm}^e} ds,$$

$$G_{nm,qp}^{012} = -\frac{\zeta}{4\pi} N_{nm} \int_S (\mathbf{n} \times \mathbf{H}_{qp}^h) \overline{\mathbf{e}_{nm}^e} ds,$$

$$G_{nm,qp}^{021} = \frac{-\zeta}{4\pi} N_{nm} \int_S (\mathbf{n} \times \mathbf{H}_{qp}^e) \overline{\mathbf{h}_{nm}^e} ds,$$
(21)

$$G_{nm,qp}^{022} = \frac{-\zeta}{4\pi} N_{nm} \int_S (\mathbf{n} \times \mathbf{H}_{qp}^h) \overline{\mathbf{h}_{nm}^e} ds.$$

The additional terms for the matrix elements in the case of a nonzero impedance are expressed in the form of the following integrals:

$$G_{nm,qp}^{z11} = \frac{-\zeta}{4\pi} N_{nm} \int_S [\mathbf{n} \times (\mathbf{n} \times \mathbf{H}_{qp}^e)]\big|_S \overline{\mathbf{h}_{nm}^e} ds,$$

$$G_{nm,qp}^{z12} = \frac{-\zeta}{4\pi} N_{nm} \int_S [\mathbf{n} \times (\mathbf{n} \times \mathbf{H}_{qp}^h)]\big|_S \overline{\mathbf{h}_{nm}^e} ds,$$
(22)

$$G_{nm,qp}^{z21} = \frac{-1}{4\pi\zeta} N_{nm} \int_S [\mathbf{n} \times (\mathbf{n} \times \mathbf{H}_{qp}^e)]\big|_S \overline{\mathbf{e}_{nm}^e} ds,$$

$$G_{nm,qp}^{z22} = \frac{-1}{4\pi\zeta} N_{nm} \int_S [\mathbf{n} \times (\mathbf{n} \times \mathbf{H}_{qp}^h)]\big|_S \overline{\mathbf{e}_{nm}^e} ds.$$

In expressions (20)–(22),

$$\mathbf{h}_{nm}^e = \frac{ik}{\zeta} \nabla \times (\mathbf{r}\chi_n^m), \quad \mathbf{e}_{nm}^e = \nabla \times \nabla \times (\mathbf{r}\chi_n^m).$$

In the spherical coordinate system,

$$\mathbf{n} = \frac{1}{\kappa}(\mathbf{i}_r\rho\sin\theta - \mathbf{i}_\theta\rho_\theta'\sin\theta - \mathbf{i}_\varphi\rho_\varphi'), \quad ds = k\rho d\theta d\varphi,$$

$$\kappa = \sqrt{(\rho^2 + (\rho_\theta')^2)\sin^2\theta + (\rho_\varphi')^2},$$

where $r = \rho(\theta, \varphi)$ is the equation of the surface S in the spherical coordinate system.

If the scatterer is a body of revolution, i.e., $\rho(\theta, \varphi) = \rho(\theta)$, system (18) takes the form

$$\begin{cases} a_{nm} = a_{nm}^0 + \sum_{q=|m|}^{\infty} (G_{nm,qm}^{11}a_{qm} + G_{nm,qm}^{12}b_{qm}) \\ b_{nm} = b_{nm}^0 + \sum_{q=|m|}^{\infty} (G_{nm,qm}^{21}a_{qm} + G_{nm,qm}^{22}b_{qm}) \\ n = 1, 2, \ldots; \quad |m| \le n, \end{cases}$$
(23)

where the quantities $G_{nm,qm}^{ij}$ are expressed through single integrals.

In order to substantiate the algorithm developed, one has to asymptotically estimate the matrix elements and the right-hand sides of system (18) for large subscripts n and q. This procedure is identical to that used in [7]. For instance, it can be shown [7] that, for $n \gg q$,

$$\left|G_{nm,qp}^{ij}\right| \le \text{const}\frac{\sigma_1^n}{nn!},$$

where

$$\sigma_1 = \max_{\theta_0^s, \varphi_0, s} \left|\frac{k\rho(\theta_0^s, \varphi_0)}{2}\exp(is\theta_0^s)\right|,$$
(24)

and θ_0^s and φ_0 are determined by

$$\frac{\rho_\theta'(\theta, \varphi)}{\rho(\theta, \varphi)}\bigg|_{\theta = \theta_0^s, \varphi_0} = -is; \quad s = \pm 1;$$

$$\frac{\rho_\varphi'(\theta, \varphi)}{\rho(\theta, \varphi)}\bigg|_{\varphi = \varphi_0, \theta_0^s} = 0; \quad \exp(is\theta_0^s) = 0.$$
(25)

Maximum (24) is looked for among those roots of system (25) falling (after the substitution $\xi = \rho(\theta, \varphi)\exp(i\theta)$) into the contours C_φ that result from mapping the section obtained when the plane $(\varphi, \varphi + \pi)$ cuts the surface S onto the complex plane $z = r\exp(i\alpha)$. These roots are the principal singularities of the extension of the diffraction field into the scatterer [9]. If the function $\rho(\theta, \varphi)$ has nonanalytical points, they have to be taken into account in calculating maximum (24).

Similarly, at $q \gg n$,

$$\left|G_{nm,qp}^{ij}\right| \le \text{const}\frac{q!}{\sigma_2^q},$$

where

$$\sigma_2 = \max_{\theta_0^s, \varphi_0, s} \left|\frac{k\rho(\theta_0^s, \varphi_0)}{2}\exp(is\theta_0^s)\right|.$$
(26)

Maximum (26) is searched for among those roots of (25) corresponding (after the substitution $\xi = \rho(\theta, \varphi)\exp(i\theta)$) to points outside the above contours C_φ on the plane $z = r\exp(i\alpha)$.

In a similar way, one can show that, for $n \gg 1$,

$$\left|a_{nm}^0\right|, \left|b_{nm}^0\right| \leq \text{const} \frac{\sigma^n}{nn!}, \quad \sigma = \max(\sigma_1, \sigma_0),$$

where $\sigma_0 = kr_0/2$ and r_0 is the distance to the farthest (from the origin) point inside S that corresponds to the singularity of the function $(\mathbf{n} \times \mathbf{H}^0)|_S(\theta, \varphi)$ in its extension to the region of complex angles θ. If $\mathbf{H}^0$ is the plane wave field, $\sigma_0 = 0$; i.e., $\sigma = \sigma_1$.

From the estimates found, it follows that one has to substitute the unknown coefficients in (18) as [6, 7]

$$a_{nm} = \frac{\sigma^n}{n!} x_{nm}, \quad b_{nm} = \frac{\sigma^n}{n!} y_{nm}.$$

Then, for example, system (23) takes the form

$$\begin{cases} x_{nm} = x_{nm}^0 + \sum_{q=|m|}^{\infty} (g_{nm,qm}^{11} x_{qm} + g_{nm,qm}^{12} y_{qm}) \\ y_{nm} = y_{nm}^0 + \sum_{q=|m|}^{\infty} (g_{nm,qm}^{21} x_{qm} + g_{nm,qm}^{22} y_{qm}) \\ n = 1, 2, \ldots; \quad |m| \leq n, \end{cases} \quad (27)$$

where

$$x_{nm}^0 = \frac{n!}{\sigma^n} a_{nm}^0, \quad y_{nm}^0 = \frac{n!}{\sigma^n} b_{nm}^0,$$

$$g_{nm,qm}^{jl} = G_{nm,qm}^{jl} \frac{n!}{q!} \sigma^{q-n}.$$

System (27), as well as the general system of equations found from (18), is solvable by the reduction method subject to

$$\sigma_2 > \sigma. \quad (28)$$

If the initial field is a plane wave, condition (28) restricts only the shape of the scatterer. It must belong to the class of weakly nonconvex bodies [7], which all convex bodies are.

RESULTS OF NUMERICAL SIMULATIONS

Consider solutions to particular scattering problems obtained by the method developed. We considered scattering by axisymmetric bodies, such as a prolate spheroid, a finite circular cylinder, a circular cylinder with an analytic boundary, and a sphere. The Z axis was taken to be the symmetry axis.

The pattern coefficients were calculated by solving a finite system of form (27) in which the upper limit of

summation is equal to N:

$$\begin{cases} x_{nm} = x_{nm}^0 + \sum_{q=|m|}^{N} (g_{nm,qm}^{11} x_{qm} + g_{nm,qm}^{12} y_{qm}) \\ y_{nm} = y_{nm}^0 + \sum_{q=|m|}^{N} (g_{nm,qm}^{21} x_{qm} + g_{nm,qm}^{22} y_{qm}) \\ n = 1, 2, \ldots; \quad |m| \leq n. \end{cases} \quad (29)$$

The solution of system (29) is reduced to the solution of $2N + 1$ smaller dimension systems of linear algebraic equations for each $m (-N \leq m \leq N)$. The minimal and maximal matrix dimensions are 2×2 and $2N \times 2N$, respectively, whereas the matrix dimension in the general case (when the symmetry of rotation is neglected) would be equal to $2(N^2 + 2N) \times 2(N^2 + 2N)$.

Let us first study the scattering patterns for perfect electrical and magnetic conductors: a prolate spheroid with $ka = 10$ and $kc = 20$ (a and b are the minor and major semiaxes, respectively) and a finite circular cylinder with $ka = 10$ and $kh = 40$ (a is the radius of the base and h is the height) into which this spheroid is
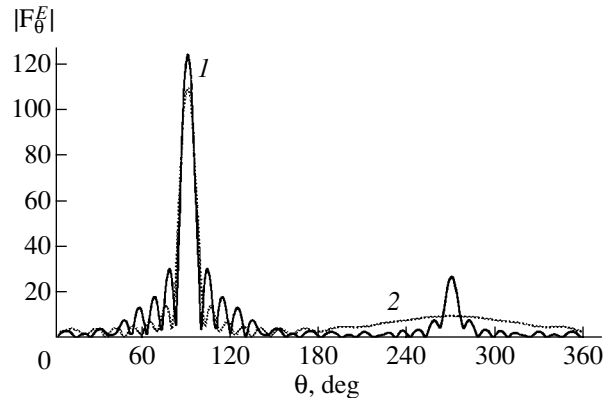

Fig. 1. Scattering patterns for perfect electrical conductors when a plane wave is incident normally to the axis of revolution: (*1*) spheroid and (*2*) cylinder.
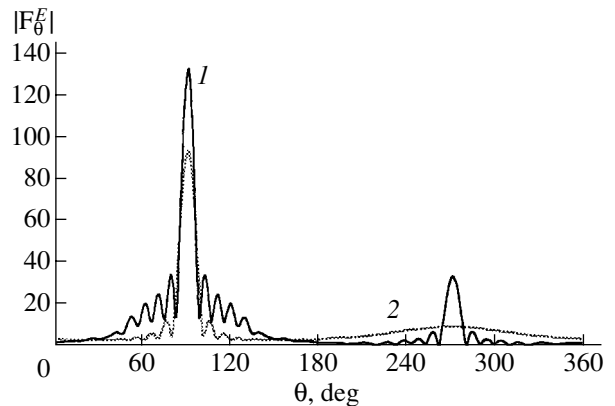

Fig. 2. The same as in Fig. 1 for perfect magnetic conductors: (*1*) spheroid and (*2*) cylinder.
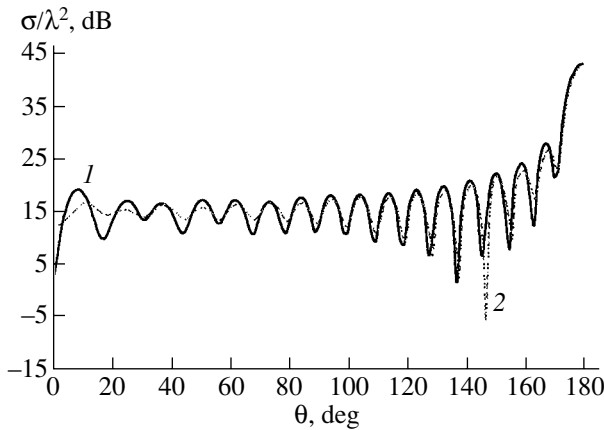
Fig. 3. Two-position scattering cross section of a conducting sphere with a radius 3λ, coating thickness $t = 0.05\lambda$, relative permittivity $\varepsilon_r = 4 - i$, and permeability $\mu_r = 1$. (*1*) PEM for $N = 25$ and $Z = -5.48 - 136.69i$ and (*2*) exact solution (Mie series) [15].

inscribed. We consider the case when a unit plane wave is incident on them at angles $\theta_0 = 90°$ and $\varphi_0 = 0°$ (normal incidence). Here, the incident field is polarized as follows:

$$\mathbf{E}^0 = \mathbf{i}_z \exp(-ikr\sin\theta\cos\varphi),$$

$$\mathbf{H}^0 = -\frac{1}{\zeta}\mathbf{i}_y \exp(-ikr\sin\theta\cos\varphi).$$

Figures 1 and 2 show the values of $|F_\theta^E|$ in the azimuthal plane $\varphi = [0, \pi]$. Curve *1* corresponds to the spheroid ($N = 40$); curve *2*, to the cylinder ($N = 44$). For a perfect electrical conductor, the impedance Z is equal to zero (Fig. 1). In the case of a perfect magnetic conductor (Fig. 2), one must put $Z = \infty$; in applications,
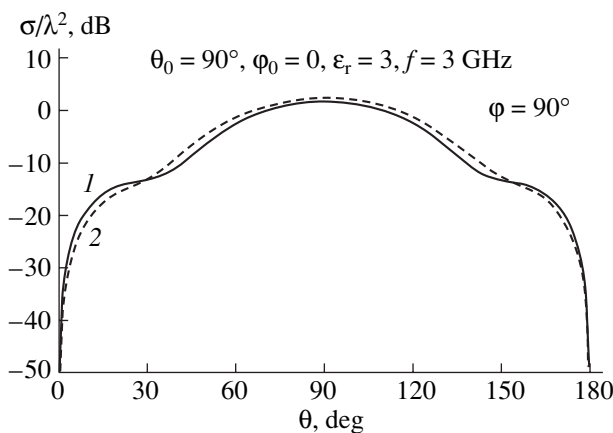
however, the impedance may be set equal to $1000\zeta_0 i$, where ζ_0 is the impedance of free space.

Now we pass on to scattering by coated bodies. An insulating coating with a thickness t, permittivity ε, and permeability μ is characterized by the impedance Z [13, 14]:

$$Z = -i\sqrt{\frac{\mu}{\varepsilon}}\tan(k_1 t),$$

where $k_1 = kn$ and n is the refractive index.

Figures 3–5 show the results of simulating the characteristics of wave scattering by bodies with an insulating coating. On the vertical axis, the quantity σ/λ^2 is plotted, where $\sigma = \lim_{r\to\infty} 4\pi r^2 |\mathbf{E}^1|^2/|\mathbf{E}^0|^2$ is the radar cross section.

Figure 3 shows the two-position cross section for a sphere of radius $a = 3\lambda$ coated by an insulator with a thickness $t = 0.05\lambda$ and material parameters $\varepsilon = (4 - i)\varepsilon_0$ and $\mu = \mu_0$, where ε_0 and μ_0 are the permittivity and permeability of free space. The dashed line in Fig. 3 is the exact solution in the form of Mie series [15], while the solid line is obtained with our impedance approximation. As can be seen, the impedance approximation is quite adequate in the case of the sphere, although the refractive index of the coating material is comparatively low.

Figure 4 shows the results of calculation of σ/λ^2 in the half-plane $\varphi = 90°$ for a finite circular cylinder of radius $a = 0.26\lambda$ and height $h = \lambda$ that is coated by an insulating layer of thickness $t = 0.02\lambda$, $\varepsilon = 3\varepsilon_0$, and $\mu = \mu_0$. Here, a plane wave is also incident normally to the axis of the cylinder ($\theta_0 = 90°$ and $\varphi_0 = 0$). The dashed line is taken from [16], where the method of current integral equations and the finite-element method were

Fig. 4. Two-position cross section in the half-plane $\varphi = 90°$ for a conducting circular cylinder with the radius $a = 0.26\lambda$, height $h = \lambda$, coating thickness $t = 0.02\lambda$, $\varepsilon_r = 3$, and $\mu_r = 1$. The plane wave is incident normally to the axis of revolution. (*1*) PEM at $N = 15$ and $Z = -48.651i$ and (*2*) integral-equation and finite-element methods [16].
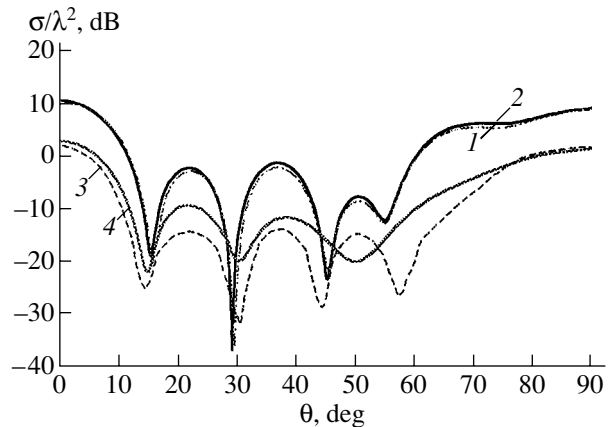

Fig. 5. One-position scattering cross section in the plane x–z for a conducting circular cylinder (radius $a = 0.5\lambda$ and height $h = 2\lambda$) uncoated and coated by a layer of thickness $t = 0.1\lambda$, $\varepsilon_r = 2 - i$, and $\mu_r = 1.5 - 0.5i$. (*1*) Uncoated cylinder (method of moments) [16], (*2*) uncoated cylinder (PEM, $Z = 0$), (*3*) coated cylinder (finite-element method) [16], and (*4*) coated cylinder (PEM, $Z = -357.134 - 326.842i$).
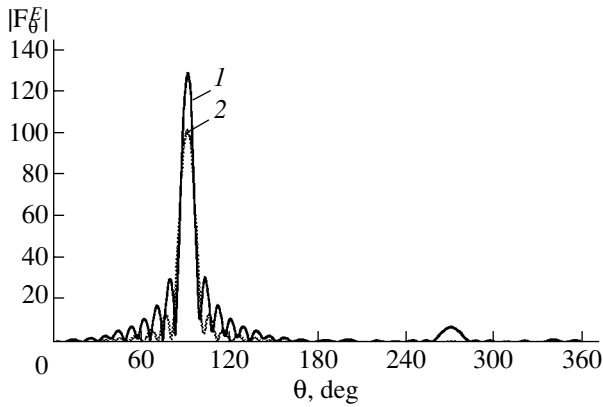
Fig. 6. Scattering patterns for black bodies when a plane wave is incident normally to the axis of revolution: (*1*) spheroid ($N = 40$) and (*2*) cylinder ($N = 44$).

Fig. 7. Scattering patterns for impedance bodies at $Z \cong -\zeta_0$ when a plane wave is incident normally to the axis of revolution: (*1*) spheroid ($N = 40$) and (*2*) cylinder ($N = 44$).

used in calculations. The results of simulation by the impedance approximation are in good agreement with those found in terms of the strict model, although here the problem is complicated by the presence of discontinuities at the boundary of the scatterer (near the discontinuities the impedance approximation is generally incorrect) and also by the fact that the dimensions of the scatterer are small compared to the wavelength.

Finally, Fig. 5 shows the one-position radar scattering cross section for a perfectly conducting finite cylinder of length $kh = 4\pi$ and radius $ka = \pi$ (curves *1*, *2*), as well as for a cylinder with an insulating coating with a thickness $t = 0.1\lambda$, $\varepsilon = (2 - i)\varepsilon_0$, and $\mu = \mu_0(1.5-0.5i)$ (curves *3*, *4*). Curves *1* and *3* are taken from [15].

In calculating the two-position cross section for the sphere shown in Fig. 3, the upper limit of summation in finite system (29) is $N = 25$. The total time taken to calculate the curve by our method is smaller than 1 min. The calculation of the corresponding curve by the finite-element method [16] involved inversion of the algebraic system with more than 187 000 unknowns and took 130 000 s of CPU time. In calculating the one-position scattering cross section for the circular cylinder (Fig. 5), $N = 23$. The total time taken to calculate each of the curves is less than 14 min. The results presented indicate the high efficiency of the approach proposed in this paper.
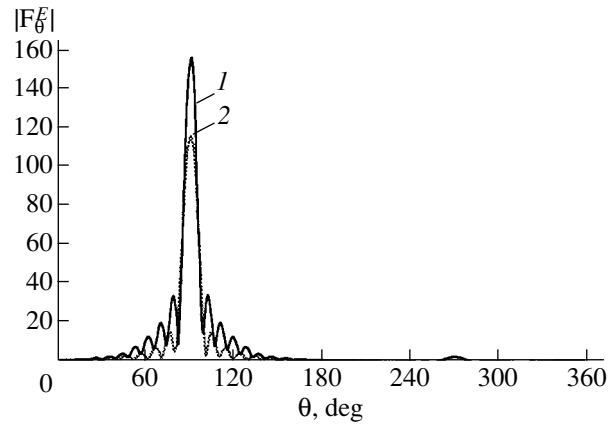
Now let us consider wave scattering by black bodies. Figure 6 plots $|F_\theta^E|$ in the plane $\varphi = [0, \pi]$ for the case when the perfectly conducting spheroid and cylinder mentioned above are black bodies. The scattering pattern in this case was calculated with the Macdonald approach [2].

According to the Sommerfeld concept [1], as was mentioned above, a metallic body with an insulating coating that has a thickness t much smaller than the body dimension and permittivity and permeability in the form

$$\varepsilon = -i\alpha\varepsilon_0, \quad \mu = -i\alpha\mu_0,$$

where $\alpha \longrightarrow \infty$, may be used as a model of a black body.

In the Sommerfeld model, the impedance is $Z \cong -\zeta_0$. Figure 7 shows $|F_\theta^E|$ for the same spheroid and cylinder in the plane $\varphi = [0, \pi]$ with the impedance corresponding to the Sommerfeld model. As can be seen, the scattering pattern for the coated body, which has the impedance $Z = -\zeta_0$, is similar to that found in the framework of the Macdonald approach. It is worth noting that the backscattering cross section in Fig. 7 is smaller than in Fig. 6 for both bodies (for the cylinder, $|F_{\theta\mathrm{Som}}^E(270°, 0)| \approx$

Table 1. Values of $\left|F_\theta^E\right|(\theta, \varphi)$ at $\theta = 120°$ and $\varphi = 0$ for the impedances $Z_0 = 0$, $Z_1 = 1000\zeta_0 i$, and $Z_2 = -\zeta_0$

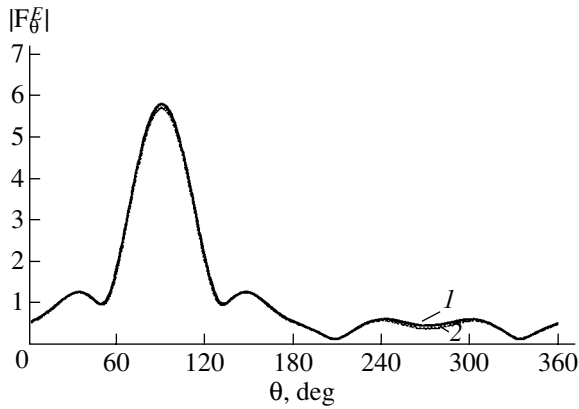| N	Spheroid			N	Cylinder		
	Z_0	Z_1	Z_2		Z_0	Z_1	Z_2
37	4.4014138	3.7177547	2.2222917	40	6.5903586	19.9426107	11.8249504
39	4.4014206	3.7177477	2.2222888	41	6.6081976	19.7724602	11.8295639
40	4.4014094	3.7177432	2.2222959	43	6.5307894	19.5571172	11.7956310
41	4.4014020	3.7177449	2.2222993	44	6.5669652	19.6462934	11.7898536

Fig. 8. Scattering patterns for black bodies of equal size when a plane wave is incident normally to the axis of revolution: (*1*) cylinder with $ka = 2$ and $kh = 8$ and (*2*) analytical cylinder.

2.6448 and $|F_{\theta\mathrm{Mac}}^{E}(270°, 0)| \approx 7.7042$; for the spheroid, $|F_{\theta\mathrm{Som}}^{E}(270°, 0)| \approx 0.0954$ and $|F_{\theta\mathrm{Mac}}^{E}(270°, 0)| \approx 0.3495$). This means that the Sommerfeld model is more appropriate for scatterers with higher absorbing properties.

The Ufimtsev theorem [3] states that the integral scattering cross section of a black body is twice as small as that of a perfectly conducting body with the same shadow contour (the boundary between the illuminated and dark regions on the body surface). This statement is valid for convex bodies for which the linear sizes and the smallest radius of curvature are much greater than the wavelength. In the case under consideration, this ratio is equal to 2.13 for the spheroid and 2.04 for the cylinder, which has a longer shadow contour.

Table 1 lists the rates of convergence of the computational algorithm for the cylinder and spheroid considered. In the case of the scatterer with the analytical surface (spheroid), four decimal places are reliably established even at $N = 40$, i.e., at $N \approx kd$, whereas only two decimal places are reliably established for $N \approx 1.1kd$ in the case of the body with the nonanalytic surface (cylinder).

The optical theorem for perfectly conducting bodies was verified. According to this theorem, the integral scattering cross section is equal to the total cross section in the direction of incident wave propagation. Table 2 presents the results of checking the optical theorem for the perfect electrical and magnetic conductors mentioned above: the spheroid and cylinder. In the case of a normally incident plane wave, the optical theorem has the form [17]

$$P_{\mathrm{S}} = P_{\mathrm{S2}} \equiv -\frac{\lambda}{\zeta}\mathrm{Im}\{F_{\theta}^{E}(\theta = 90°, \varphi = 0)\}.$$

In terms of the Kirchhoff approximation [2], the integral cross section for a black body is estimated as

$$P_{\mathrm{S}} \equiv \frac{1}{2\zeta}\int_{0}^{2\pi}\int_{0}^{\pi}|\mathbf{F}^{E}(\theta, \varphi)|^{2}\sin\theta d\theta d\varphi = P_{\mathrm{S3}} \equiv \frac{|\mathbf{E}^{0}|^{2}S_{0}}{2\zeta_{0}},$$

where S_{0} is the cross-sectional area of the shadow region.

Table 3 lists the integral scattering cross sections (P_{S}) of a black spheroid and a black cylinder that are found based on the Macdonald concept (P_{SMac}), Sommerfeld concept (P_{SSom}), and Kirchhoff approximation (P_{S3}).

Finally, Fig. 8 plots the values of $|F_{\theta}^{E}|$ in the plane $\varphi = [0, \pi]$ for the following black bodies of equal size: a finite circular cylinder with $ka = 2$ and $kh = 8$ and a finite circular cylinder with an analytical boundary for $N = 19$ in system (29). The canonical equation of an analytical cylinder has the form

$$(x^{2} - a^{2})(z^{2} - (h/2)^{2}) = \varepsilon^{2}, \quad \varepsilon \leq 0.1ka,$$

where $\varepsilon = 0.01$ is the radius of curvature at the end of the cylinder.

The scattering patterns for both bodies in Fig. 8 coincide within the graphical accuracy.

CONCLUSIONS

From the analysis of scattering by impedance bodies with irregular boundaries, one can conclude that the pattern equation method can effectively solve problems of such a kind. The PEM-based computational algorithm is very fast and does not require elimination of current singularities at joints of boundary segments with different curvatures. The impedance approximation is applicable to simulating the characteristics of scatterers with insulating coatings and weakly reflecting scatterers, including those having surface irregularities and low-optical-density coatings.

Table 2. Optical theorem for the impedances $Z_0 = 0$ and $Z_1 = 1000\zeta_0 i$

N	Spheroid		Cylinder	
	Z_0	Z_1	Z_0	Z_1
P_{S}	1.8322582	1.5556632	2.0700647	2.2404235
P_{S2}	1.8322582	1.5556631	2.0874581	2.2232507

Table 3. Integral scattering cross sections for a black spheroid and a black cylinder

	Spheroid	Cylinder
P_{SMac}	0.794916	1.035255
P_{SSom}	0.940357	1.307298
P_{S3}	0.833333	1.061039

REFERENCES

1. A. Sommerfeld, *Optics* (Academic, New York, 1954; Inostrannaya Literatura, Moscow, 1953).

2. L. N. Zakhar'ev and A. A. Lemanskiĭ, *Scattering of Waves by Blackbodies* (Sov. Radio, Moscow, 1972).

3. P. Ya. Ufimtsev, *Three Lectures on the Physical Theory of Diffraction* (Leningr. Gos. Univ., Leningrad, 1972).

4. S. K. Sharma and A. R. Jones, J. Quant. Spectrosc. Radiat. Transf. **79–80**, 1051 (2003).

5. M. A. Leontovich, *Propagation of Radio Waves* (Akad. Nauk SSSR, Moscow, 1948).

6. A. G. Kyurkchan, Dokl. Akad. Nauk **325**, 273 (1992) [Sov. Phys. Dokl. **37**, 338 (1992)].

7. A. G. Kyurkchan, Dokl. Akad. Nauk **337**, 728 (1994) [Phys. Dokl. **39**, 546 (1994)].

8. A. G. Kyurkchan, Dokl. Akad. Nauk **322**, 686 (1992) [Sov. Phys. Dokl. **37**, 70 (1992)].

9. A. G. Kyurkchan, B. Yu. Sternin, and V. E. Shatalov, Usp. Fiz. Nauk **166**, 1285 (1996) [Phys. Usp. **39**, 1221 (1996)].

10. Do Dyk Tkhang and A. G. Kyurkchan, Akust. Zh. **49**, 55 (2003) [Acoust. Phys. **49**, 43 (2003)].

11. A. G. Kyurkchan, Radiotekh. Élektron. (Moscow) **45**, 1078 (2000).

12. C. H. Papas, *Theory of Electromagnetic Wave Propagation* (McGraw-Hill, New York, 1965).

13. Ya. N. Fel'd and L. S. Benenson, *Antenna Feeders* (Voenno-Vozdushn. Inzh. Akad. im. N. E. Zhukovskogo, Moscow, 1959), Part 2.

14. Do-Hoon Kwon, R. J. Burkholder, and P. H. Pathak, IEEE Trans. Antennas Propag. **49**, 583 (2001).

15. Xin-Qing Sheng, Jian-Ming Jin, Jiming Song, *et al.*, IEEE Trans. Antennas Propag. **46**, 583 (1998).

16. T. Cwik, C. Zuffada, and V. Jamnejad, IEEE Trans. Antennas Propag. **44**, 453 (1996).

17. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983; Mir, Moscow, 1986).

Translated by M. Fofanov

Method of Integral Equations in the Theory of Microstructured Optical Fibers

A. B. Sotsky and L. I. Sotskaya

Institute of Applied Optics, National Academy of Sciences of Belarus, Mogilev, 212793 Belarus
e-mail: ipo@physics.belpak.mogilev.by
Received July 8, 2003

Abstract—A method for designing microstructured optical fibers that is based on exact integral equations for the transverse components of the magnetic field of the mode is proposed. A solution to the vector waveguide problem for fibers with a finite number of circular capillaries in the round cavity of the cladding can be refined successively. Quartz fibers with hexagonal capillary rings are also studied. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Microstructured optical fibers, the guiding regions of which are formed by microscopic (mostly air) channels in the insulating environment, have recently become the object of intensive research, because they offer a unique combination of dispersion, polarization, and nonlinear properties [1–8]. A number of theoretical methods to analyze modes propagating in microstructured fibers has been suggested [9–15]. In [9–15], the cross section of the fiber was simulated by a set of circular inclusions embedded in a homogeneous medium that either occupies the entire cross section or is surrounded by a cladding with different permittivity. The recently elaborated more general and physically more adequate model [14, 15] was investigated by the numerical multipole method. This method is also of interest because it takes into account the vector character of the waveguide problem and makes it possible to describe the mode leakage effect. The multipole method uses the expansion of the longitudinal components of the electromagnetic field in cylindrical harmonics whose amplitudes are found from the continuity conditions for the tangential components at the inclusion–cladding interfaces. However, the associated computational scheme is very tedious [14, 15]. In this work, we suggest an alternative method for calculating the eigenmodes and leaky modes of microstructured optical fibers. It is based on exact integral equations for the transverse components of the magnetic field and results in relationships that are an analogue of the Ewald–Oseen extinction theorem [16]. Unlike the multipole method, our method is simpler in mathematical statement (there is no need to represent interinclusion fields) and allows one to study fibers made of a set of capillaries (capillary fibers). Circular fibers considered previously is a specific case of such structures, which are of practical interest [4, 5, 12]. Below, we justify the method of integral equations, illustrate the convergence of solutions, and compare the results obtained with this method and with the multipole method for fibers with circular inclusions. The dispersion characteristics and mode fields of quartz capillary fibers are studied. Conditions for the maximal contraction of the fundamental mode field are determined. Capillary fibers are shown to combine a high spatial confinement of the mode energy, zero dispersion of the group velocity, and low leakage losses.

MATHEMATICAL STATEMENT OF THE METHOD

Consider a fiber consisting of n circular capillaries distributed over the circular cladding (Fig. 1a). The cladding has a radius A, and the inner and outer radii of an lth capillary are a_l and b_l, respectively. We use the global Cartesian system (x, y) where the origin $(0, 0)$ coincides with the center of the cladding and the point (x_l, y_l) is the center of the lth capillary, as well as global polar coordinates (ρ_l, φ_l) where $\rho\lambda_l = 0$ corresponds to the center of the lth capillary. It is assumed that the domains $\rho > A$, $\rho_l < a_l$, and $a_l < \rho_l < b_l$ are occupied by media with respective permittivities ε_c, $\varepsilon_l^{(1)}$, and $\varepsilon_l^{(2)}$ and that the space between the capillaries in the domain $\rho < A$ is occupied by a medium with a permittivity ε_s. The permittivities ε_c, $\varepsilon_l^{(1)}$, $\varepsilon_l^{(2)}$, and ε_s may be complex quantities. A specific case of such fibers is that shown in Fig. 1b. A homogeneous rod at the center will be considered as a capillary with the number $l = 1$ ($x_l = y_l = 0$) and $\varepsilon_l^{(1)} = \varepsilon_l^{(2)}$. This fiber will be discussed in greater detail below.

Let the dependence of the optical field on time t and longitudinal coordinate z involve the factor $\exp[i(\omega t - \beta z)]$, where β is the complex mode propagation constant. Then, the transverse components of the magnetic

field H_j ($j = x, y$) of the mode will satisfy Eqs. (5) in [17]. With the Hankel function represented in the form

$$H_0^{(2)}(\chi_s r) = \frac{1}{\pi}\int\limits_{\Gamma}\frac{dk}{\sqrt{\chi_s^2 - k^2}}$$

$$\times \exp[-ik(x'-x) - i\sqrt{\chi_s^2 - k^2}|y'-y|],$$

we can write

$$H_j(x, y) = \frac{i}{4}\int\limits_{-\infty}^{\infty} dx' \int\limits_{-\infty}^{\infty} H_0^{(2)}(\chi_s r)f_j(x', y')dy'. \qquad (1)$$

Here, $\chi_s = \sqrt{k_0^2\varepsilon_s - \beta^2}$ ($\mathrm{Re}\chi_s > 0$ if $\mathrm{Re}(k_0^2\varepsilon_s - \beta^2) > 0$, $\mathrm{Im}\chi_s < 0$ if $\mathrm{Re}(k_0^2\varepsilon_s - \beta^2) < 0$), $k_0 = \lambda_0^{-1}2\pi$ is the wavenumber of free space, $r = \sqrt{(x'-x)^2 + (y'-y)^2}$, the contour Γ is defined in [17],

$$f_j(x, y) = E\frac{\partial\varepsilon}{\partial\xi_j} - k_0^2\Delta\varepsilon H_j, \qquad (2)$$

$$E = \varepsilon^{-1}(\nabla_x H_y - \nabla_y H_x), \qquad (3)$$

$\xi_x = -y$, $\xi_y = x$, $\Delta\varepsilon = \varepsilon(x, y) - \varepsilon_s$, and $\varepsilon(x, y)$ is the permittivity of the medium.

Since the functions $f_j(x, y)$ are other than zero only in the capillaries and cladding, relationships (1) within these domains are exact integral equations. Outside these domains, they are formulas for direct calculation.

To algebraize Eqs. (1), we should define bases to represent the functions $H_j(x, y)$. Inside the capillaries, these functions must be finite at $\rho_l \longrightarrow 0$ and satisfy the Helmholtz equations

$$[\nabla_x^2 + \nabla_y^2 + (\chi_l^{(k)})^2]H_j = 0, \qquad (4)$$

where $k = 1$ at $\rho_l < a_l$, $k = 2$ at $a_l < \rho_l < b_l$, and $(\chi_l^{(k)})^2 = k_0^2\varepsilon_l^{(k)} - \beta^2$.

Thus, we may write

$$H_j = \sum_{\nu = -\infty}^{\infty} Z_{j\nu}^{(k)}(\rho_l)\exp(i\nu\varphi_l), \qquad (5)$$

where $k = 1$ at $\rho_l < a_l$, $k = 2$ at $a_l < \rho_l < b_l$,

$$Z_{j\nu}^{(1)}(\rho_l) = A_{l\nu}^{(j)}J_\nu(\chi_l^{(1)}\rho_l),$$

$$Z_{j\nu}^{(2)}(\rho_l) = B_{l\nu}^{j1}J_\nu(\chi_l^{(2)}\rho_l) + B_{l\nu}^{j2}H_\nu^{(2)}(\chi_l^{(2)}\rho_l),$$

$J_\nu(\chi_l^{(s)}\rho_l)$ are the Bessel functions, and $A_{l\nu}^{(j)}$ and $B_{l\nu}^{jk}$ are unknown coefficients.

In the domain $\rho > A$, the functions $H_j(x, y)$ must satisfy the Helmholtz equation
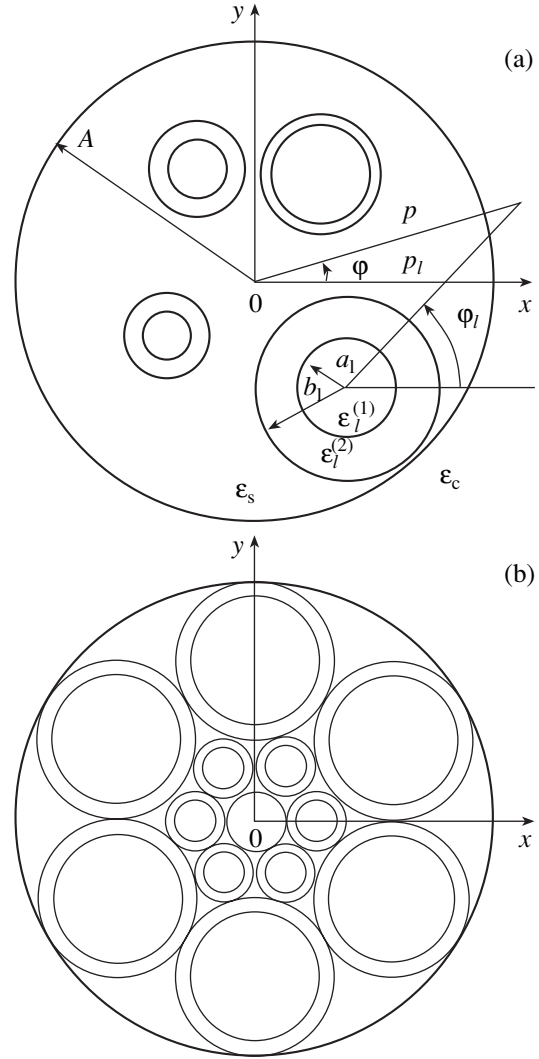
$$(\nabla_x^2 + \nabla_y^2 + \chi_c^2)H_j = 0$$

Fig. 1. Cross sections of the fibers.

and, at $\rho \longrightarrow \infty$, have the asymptotics

$$H_j \sim \Phi(\varphi)(\sqrt{\rho})^{-1}\exp(-i\chi_c\rho), \qquad (6)$$

where $\Phi(\varphi)$ is a continuous function of angular variable and $\chi_c = \sqrt{k_0^2\varepsilon_c - \beta^2}$ ($\mathrm{Im}\chi_c < 0$ for the eigenmodes and $\mathrm{Re}\chi_c > 0$ for leaky modes) [17].

Hence, with $\rho > A$, we have

$$H_j = \sum_{\nu = -\infty}^{\infty} Z_{j\nu}^{(c)}(\rho)\exp(i\nu\varphi), \qquad (7)$$

where $Z_{j\nu}^{(c)}(\rho) = C_\nu^{(j)}H_\nu^{(2)}(\chi_c\rho)$ and $C_\nu^{(j)}$ are unknown coefficients.

Note that, by virtue of asymptotics (6) and $H_0^{(2)}(\chi_c r) \sim (\sqrt{\rho})^{-1}\exp(-i\chi_s\rho)$ ($\rho \longrightarrow \infty$), the rights of Eqs. (1) exist if $\varepsilon_s = \varepsilon_c$ or $\mathrm{Im}(\chi_s + \chi_c) < 0$. This condition

is usually met in the cases of interest. Recast Eqs. (1) in the form

$$H_j(x, y) = \Omega_j^{(c)}(x, y) + \sum_{l=1}^{n} \Omega_j^{(l)}(x, y), \qquad (8)$$

where the terms $\Omega_j^{(c)}(x, y)$ and $\Omega_j^{(l)}(x, y)$ are the contributions to the rights of (1) from the cladding and lth capillary, respectively.

To define these terms in explicit form, we note that differentiating the step function $\varepsilon(x, y)$ in (2) yields Dirac delta functions. Integrating these functions yields

$$\Omega_j^{(c)}(x, y)$$
$$= \frac{iA}{4}(\varepsilon_s - \varepsilon_c) \int_0^{2\pi} P_j^{(c)}[EH_0^{(2)}(\chi_c r)]_{\rho' = A} d\varphi' + I_c, \qquad (9)$$

$$\Omega_j^{(l)}(x, y) = \frac{i}{4} \int_0^{2\pi} P_j^{(l)}\{b_l(\varepsilon_l^{(2)} - \varepsilon_s)[EH_0^{(2)}(\chi_c r)]_{\rho'_l = b_l}$$
$$- a_l(\varepsilon_l^{(2)} - \varepsilon_l^{(1)})[EH_0^{(2)}(\chi_c r)]_{\rho'_l = a_l}\} d\varphi' + I_1 + I_2. \qquad (10)$$

Here,

$$I_c = \frac{ik_0^2}{4}(\varepsilon_s - \varepsilon_c) \int\int_{\rho' > A} H_0^{(2)}(\chi_s r) H_j(x', y') dx' dy',$$

$$I_1 = \frac{ik_0^2}{4}(\varepsilon_s - \varepsilon_l^{(1)}) \int\int_{\rho'_l < a_l} H_0^{(2)}(\chi_s r) H_j(x', y') dx' dy',$$

$$I_2 = \frac{ik_0^2}{4}(\varepsilon_s - \varepsilon_l^{(2)}) \int\int_{a_l < \rho'_l < b_l} H_0^{(2)}(\chi_s r) H_j(x', y') dx' dy', \qquad (11)$$

$$P_x^{(l)} = \sin\varphi'_l, \quad P_y^{(l)} = -\cos\varphi'_l, \quad P_x^{(c)} = \sin\varphi',$$
$$P_y^{(c)} = -\cos\varphi';$$
$$x = \rho\cos\varphi, \quad y = \rho\sin\varphi, \quad x' = \rho'\cos\varphi',$$
$$y' = \rho'\sin\varphi'$$

in expression (9);

$$x = x_l + \rho_l\cos\varphi_l, \quad y = y_l + \rho_l\sin\varphi_l,$$
$$x' = x_l + \rho'_l\cos\varphi'_l, \quad y' = y_l + \rho'_l\sin\varphi'_l$$

in expression (10); and the function $E(x, y)$ is calculated from (3), (5), and (7). To simplify the integral I_1, we introduce continuous functions $h_{jl}^{(1)}(x, y)$ that satisfy the Helmholtz equation

$$(\nabla_x^2 + \nabla_y^2 + \chi_s^2)h_{jl}^{(1)} = 0 \qquad (12)$$

in the domain $\rho_l > a_l$ and have the form

$$h_{jl}^{(1)} = \begin{cases} H_j & \text{for} \quad \rho_l < a_l \\ \sum_{\nu = -\infty}^{\infty} Z_{j\nu}^{(1)}(a_l)[H_\nu^{(2)}(\chi_s a_l)]^{-1} \\ \times H_\nu^{(2)}(\chi_s\rho_l)\exp(i\nu\varphi_l) & \text{for} \quad \rho_l > a_l. \end{cases} \qquad (13)$$

Let us also take advantage of the equation [18]

$$(\nabla_x^2 + \nabla_y^2 + \chi_s^2)H_0^{(2)}(\chi_s r) = -4i\delta(x' - x)\delta(y' - y), (14)$$

the right of which involves Dirac functions. Multiplying Eqs. (4), (12), and (14) by $H_0^{(2)}(\chi_s r)$, $H_0^{(2)}(\chi_s r)$, and $h_{jl}^{(1)}(x', y')$, respectively; integrating the products; and applying Green's theorem [19] to the resultant expressions, we arrive at

$$I_1 = h_{jl}^{(1)} + \frac{ia_l}{4} \int_0^{2\pi} [H_0^{(2)}(\chi_s r)]_{\rho'_l = a_l} \frac{\partial}{\partial\rho'_l} h_{jl}^{(1)} \Big|_{\rho'_l = a_l + 0}^{\rho'_l = a_l - 0} d\varphi'_l. \qquad (15)$$

Similar transformations applied to the integrals I_2 and I_c yield

$$I_2 = h_{jl}^{(2)} + \frac{ib_l}{4} \int_0^{2\pi} [H_0^{(2)}(\chi_s r)]_{\rho'_l = b_l} \frac{\partial}{\partial\rho'_l} h_{jl}^{(2)} \Big|_{\rho'_l = b_l + 0}^{\rho'_l = b_l - 0} d\varphi'_l$$
$$- \frac{ia_l}{4} \int_0^{2\pi} [H_0^{(2)}(\chi_s r)]_{\rho'_l = a_l} \frac{\partial}{\partial\rho'_l} h_{jl}^{(2)} \Big|_{\rho'_l = a_l - 0}^{\rho'_l = a_l + 0} d\varphi'_l, \qquad (16)$$

$$I_c = h_{jc} - \frac{iA}{4} \int_0^{2\pi} [H_0^{(2)}(\chi_s r)]_{\rho' = A} \frac{\partial}{\partial\rho} h_{jc} \Big|_{\rho' = A - 0}^{\rho' = A + 0} d\varphi', \qquad (17)$$

$$h_{jl}^{(2)} = \begin{cases} \sum_{\nu = -\infty}^{\infty} Z_{j\nu}^{(2)}(b_l)[H_\nu^{(2)}(\chi_s b_l)]^{-1} \\ \times H_\nu^{(2)}(\chi_s\rho_l)\exp(i\nu\varphi_l) & \text{for} \quad \rho_l > b_l, \\ H_j & \text{at} \quad a_l < \rho_l < b_l, \\ \sum_{\nu = -\infty}^{\infty} Z_{j\nu}^{(2)}(a_l)[J_\nu(\chi_s a_l)]^{-1} \\ \times J_\nu(\chi_s\rho_l)\exp(i\nu\varphi_l) & \text{for} \quad \rho_l < a_l, \end{cases} \qquad (18)$$

$$h_{jc} = \begin{cases} H_j & \text{for} \quad \rho > A, \\ \sum_{\nu = -\infty}^{\infty} Z_{j\nu}^{(c)}(A)[J_\nu(\chi_s A)]^{-1} \\ \times J_\nu(\chi_s\rho)\exp(i\nu\varphi) & \text{for} \quad \rho < A. \end{cases} \qquad (19)$$

In view of (3), (5), (7), (13), (18), and (19), as well as of the identity [20]

$$H_0^{(2)}(\chi_s r) = \sum_{\nu=-\infty}^{\infty} \exp[i\nu(\varphi - \varphi')]$$

$$\times \begin{cases} H_\nu^{(2)}(\chi_s\rho')J_\nu(\chi_s\rho) & \text{for} \quad \rho < \rho' \\ J_\nu(\chi_s\rho')H_\nu^2(\chi_s\rho) & \text{for} \quad \rho > \rho', \end{cases} \quad (20)$$

which was written subject to (11), the integrals in (9), (10), and (15)–(17) can be taken analytically. As a result, the rights of (8) are represented through series in cylindrical functions defined in the local coordinates ρ_l and φ_l ($l = 1, 2, \ldots, n$) and in the global coordinates ρ and φ. Reducing the series to unique coordinate systems and applying Graf's addition theorem for cylindrical functions [20], we transform Eqs. (8) into the form

$$H_x + (-1)^p iH_y = H_x + (-1)^p iH_y + \delta H_{lp}^{(1)} \quad \text{for} \quad (21)$$
$$\sqrt{(x-x_l)^2 + (y-y_l)^2} < a_l,$$

$$H_x + (-1)^p iH_y = H_x + (-1)^p iH_y + \delta H_{lp}^{(2)} \quad \text{for} \quad (22)$$
$$a_l < \sqrt{(x-x_l)^2 + (y-y_l)^2} < b_l,$$

$$H_x + (-1)^p iH_y = H_x + (-1)^p iH_y + \delta H_p \quad \text{for} \quad (23)$$
$$\sqrt{x^2 + y^2} > A.$$

Here, ρ takes the values 0 or 1,

$$\delta H_{lp}^{(1)} = \sum_{\nu=-\infty}^{\infty} (U_{l\nu}^{1p} + V_{l\nu}^{(p)})J_\nu(\chi_s\rho_l)\exp(i\nu\varphi_l), \quad (24)$$

$$\delta H_{lp}^{(2)} = \sum_{\nu=-\infty}^{\infty} [U_{l\nu}^{2p}H_\nu^{(2)}(\chi_s\rho_l) + V_{l\nu}^{(p)}J_\nu(\chi_s\rho_l)] \quad (25)$$
$$\times \exp(i\nu\varphi_l),$$

$$\delta H_p = \sum_{\nu=-\infty}^{\infty} W_\nu^{(p)}H_\nu^{(2)}(\chi_s\rho)\exp(i\nu\varphi), \quad (26)$$

and the quantities $U_{l\nu}^{kp}$, $V_{l\nu}^{(p)}$, and $W_\nu^{(p)}$ are linear functions of the coefficients in series (5) and (7). The functions $U_{l\nu}^{kp}$, $V_{l\nu}^{(p)}$, and $W_\nu^{(p)}$ are easy to derive in explicit form; however, the associated expressions are very awkward. Therefore, we will give only the final results of our analysis. According to (21)–(26), we have

$$U_{l\nu}^{1p} = 0, \quad U_{l\nu}^{2p} = 0, \quad (27)$$

$$V_{l\nu}^{(p)} = 0, \quad W_\nu^{(p)} = 0, \quad (28)$$

where $l = 1, 2, \ldots, n$; $p = 0, 1$; and $\nu = 0, \pm 1, \ldots$.

One can show that Eqs. (27) turn into identities if

$$\bar{B}_{l\nu}^{pk} = Q_{pl\nu}^{k11}\bar{A}_{l\nu}^{(p)} + Q_{pl\nu}^{k21}\bar{A}_{l\nu+\sigma}^{(q)} \quad (29)$$

(where $q = 1 - p$ and $\sigma = 2(p - q)$),

$$\bar{B}_{l\nu}^{pk} = B_{l\nu}^{xk} + (-1)^p iB_{l\nu}^{yk}, \quad \bar{A}_{l\nu}^{(p)} = A_{l\nu}^{(x)} + (-1)^p iA_{l\nu}^{(y)},$$

$$Q_{0l\nu}^{111} = K_l^{(1)}\{\chi_l^{(2)}(\chi_l^{(1)})^{-1}J_\nu(z_l^{(1)})$$
$$\times [H_{\nu-1}^{(2)}(z_l^2) - H_{\nu+1}^{(2)}(z_l^{(2)})]$$
$$- H_\nu^{(2)}(z_l^{(2)})[\varepsilon_l^{(2)}(\varepsilon_l^{(1)})^{-1}J_{\nu-1}(z_l^{(1)}) - J_{\nu+1}(z_l^{(1)})]\},$$

$$Q_{0l\nu}^{121} = K_l^{(1)}[1 - \varepsilon_l^{(2)}(\varepsilon_l^{(1)})^{-1}]J_{\nu-1}(z_l^{(1)})H_\nu^{(2)}(z_l^{(2)}),$$

$$Q_{0l\nu}^{211} = K_l^{(1)}\{J_\nu(z_l^{(2)})[\varepsilon_l^{(2)}(\varepsilon_l^{(1)})^{-1}J_{\nu-1}(z_l^{(1)}) - J_{\nu+1}(z_l^{(1)})]$$
$$- \chi_l^{(2)}(\chi_l^{(1)})^{-1}J_\nu(z_l^{(1)})[J_{\nu-1}(z_l^{(2)}) - J_{\nu+1}(z_l^{(2)})]\},$$

$$Q_{0l\nu}^{221} = -K_l^{(1)}[1 - \varepsilon_l^{(2)}(\varepsilon_l^{(1)})^{-1}]J_{\nu-1}(z_l^{(1)})J_\nu^{(2)}(z_l^{(2)}),$$

$$Q_{1l\nu}^{1k1} = Q_{0l\nu}^{1k1}$$
$$- 0.5i\pi\nu[1 - \varepsilon_l^{(2)}(\varepsilon_l^{(1)})^{-1}]J_\nu(z_l^{(1)})H_\nu^{(2)}(z_l^{(2)}),$$

$$Q_{1l\nu}^{2k1} = Q_{0l\nu}^{2k1}$$
$$+ 0.5i\pi\nu[1 - \varepsilon_l^{(2)}(\varepsilon_l^{(1)})^{-1}]J_\nu(z_l^{(1)})J_\nu(z_l^{(2)}),$$

$$K_l^{(1)} = 0.25i\pi z_l^{(1)}, \quad z_l^{(1)} = \chi_l^{(1)}a_l, \quad z_l^{(2)} = \chi_l^{(2)}a_l.$$

Subject to (29), the lefts of Eqs. (28) take the form

$$V_{l\nu}^{(p)} = R_{l\nu}^{1p}\bar{A}_{l\nu}^{(p)} + R_{l\nu}^{2p}\bar{A}_{l\nu+\sigma}^{(q)}$$

$$+ \sum_{\mu=-\infty}^{\infty}\left[F_{\mu-\nu}^{(l)}T_{l\mu}^{(p)}\bar{C}_\mu^{(p)} + F_{\mu-\nu-\sigma}^{(l)}T_{2\mu-\sigma}^{(p)}\bar{C}_\mu^{(q)}\right.$$

$$\left.+ \sum_{k\neq l}(G_{\mu-\nu}^{lk}S_{k\mu}^{1p}\bar{A}_{k\mu}^{(P)} + G_{\mu-\nu-\sigma}^{lk}S_{k\mu-\sigma}^{2p}\bar{A}_{k\mu}^{(q)})\right], \quad (30)$$

$$W_\nu^{(p)} = L_{1\nu}^{(p)}\bar{C}_\nu^{(p)} + L_{2\nu}^{(p)}\bar{C}_{\nu+\sigma}^{(q)}$$

$$+ \sum_{\mu=-\infty}^{\infty}(-1)^{\mu-\nu}\sum_{l=1}^{n}(F_{\mu-\nu}^{(l)}S_{l\mu}^{1p}\bar{A}_{l\mu}^{(p)} + F_{\mu-\nu-\sigma}^{(l)}S_{l\mu-\sigma}^{2p}\bar{A}_{l\mu}^{(q)}). \quad (31)$$

Here,

$$\bar{C}_\nu^{(p)} = C_\nu^{(x)} + (-1)^p iC_\nu^{(y)},$$

$$R_{l\nu}^{1p} = Q_{pl\nu}^{111}Q_{pl\nu}^{112} + Q_{pl\nu}^{211}Q_{pl\nu}^{212} + Q_{ql\nu+\sigma}^{121}Q_{pl\nu}^{122} + Q_{ql\nu+\sigma}^{221}Q_{pl\nu}^{222},$$

$$R_{l\nu}^{2p} = Q_{pl\nu}^{121}Q_{pl\nu}^{112} + Q_{pl\nu}^{221}Q_{pl\nu}^{212} + Q_{ql\nu+\sigma}^{111}Q_{pl\nu}^{122} + Q_{ql\nu+\sigma}^{211}Q_{pl\nu}^{222},$$

$$T_{k\nu}^{(p)} = [J_\nu(z_s)]^{-1}[L_{k\nu}^{(p)}H_\nu^{(2)}(z_s) + (2 - k)H_\nu^{(2)}(z_c)],$$

$$S_{l\nu}^{kp} = [H_\nu^{(2)}(z_l^4)]^{-1}[Q_{pl\nu}^{1k1}J_\nu(z_l^{(3)})$$

$$+ Q_{pl\nu}^{2k1}H_\nu^{(2)}(z_l^{(3)}) + R_{l\nu}^{kp}J_\nu(z_l^{(4)})],$$

$$L_{1\nu}^{(0)} = K_c\{J_\nu(z_s)[H_{\nu+1}^{(2)}(z_c) - \varepsilon_s\varepsilon_c^{-1}H_{\nu-1}^{(2)}(z_c)]$$

$$+ \chi_s\chi_c^{-1}H_\nu^{(2)}(z_c)[J_{\nu-1}(z_s) - J_{\nu+1}(z_s)]\},$$

$$L_{2\nu}^{(0)} = K_c(1 - \varepsilon_s\varepsilon_c^{-1})J_\nu(z_s)H_{\nu-1}^{(2)}(z_c),$$

$$L_{k\nu}^{(1)} = L_{k\nu}^{(0)} - 0.5i\pi\nu(1 - \varepsilon_s\varepsilon_c^{-1})J_\nu(z_s)H_\nu^{(2)}(z_c),$$

$$Q_{0l\nu}^{112} = K_l^{(2)}\{H_\nu^{(2)}(z_l^{(4)})[\varepsilon_s(\varepsilon_l^{(2)})^{-1}J_{\nu-1}(z_l^{(3)}) - J_{\nu+1}(z_l^{(3)})]$$

$$- \chi_s(\chi_l^{(2)})^{-1}J_\nu(z_l^{(3)})[H_{\nu-1}^{(2)}(z_l^{(4)}) - H_{\nu+1}^{(2)}(z_l^{(4)})]\},$$

$$Q_{0l\nu}^{122} = -K_l^{(2)}[1 - \varepsilon_s(\varepsilon_l^{(2)})^{-1}]J_{\nu-1}(z_l^{(3)})H_\nu^{(2)}(z_l^{(4)}),$$

$$Q_{0l\nu}^{212} = K_l^{(2)}\{H_\nu^{(2)}(z_l^{(4)})$$

$$\times [\varepsilon_s(\varepsilon_l^{(2)})^{-1}H_{\nu-1}^{(2)}(z_l^{(3)}) - H_{\nu+1}^{(2)}(z_l^{(3)})]$$

$$- \chi_s(\chi_l^{(2)})^{-1}H_\nu^{(2)}(z_l^{(3)})[H_{\nu-1}^{(2)}(z_l^{(4)}) - H_{\nu+1}^{(2)}(z_l^{(4)})]\},$$

$$Q_{0l\nu}^{222} = -K_l^{(2)}[1 - \varepsilon_s(\varepsilon_l^{(2)})^{-1}]H_{\nu-1}^{(2)}(z_l^{(3)})H_\nu^{(2)}(z_l^{(4)}),$$

$$Q_{1l\nu}^{1k2} = Q_{0l\nu}^{1k2} + 0.5i\pi\nu[1 - \varepsilon_s(\varepsilon_l^{(2)})^{-1}]J_\nu(z_l^{(3)})H_\nu^{(2)}(z_l^{(4)}),$$

$$Q_{1l\nu}^{2k2} = Q_{0l\nu}^{2k2}$$

$$+ 0.5i\pi\nu[1 - \varepsilon_s(\varepsilon_l^{(2)})^{-1}]H_\nu^{(2)}(z_l^{(3)})H_\nu^{(2)}(z_l^{(4)}),$$

$$K_l^{(2)} = 0.25i\pi z_l^{(3)}, \quad K_c = 0.25i\pi z_c,$$

$$z_l^{(3)} = \chi_l^{(2)}b_l, \quad z_l^{(4)} = \chi_s b_l, \quad z_s = \chi_s A, \quad z_c = \chi_c A,$$

$$F_\nu^{(l)} = J_\nu(\chi_s\rho^{(l)})\exp(i\nu\varphi^{(l)}),$$

$$G_\nu^{lk} = H_{-\nu}^{(2)}(\chi_s\rho_{lk})\exp(i\nu\varphi_{lk}),$$

$\rho^{(l)}$ and $\varphi^{(l)}$ are the global polar coordinates of the center of an lth capillary, and ρ_{lk} and φ_{lk} are the polar coordinates of the center of a kth capillary in the local coordinate system related to the lth capillary.

As follows from (30) and (31), Eqs. (28) produce a homogeneous algebraic system that has the matrix representation

$$MX = 0, \tag{32}$$

where X is the columnar vector composed of the unknown coefficients $\bar{A}_{l\nu}^{(p)}$ and $\bar{C}_\nu^{(p)}$.

Note that the fulfillment of Eqs. (27) and (28) means that waves satisfying the Helmholtz equation

$$(\nabla_x^2 + \nabla_y^2 + \chi_s^2)\psi = 0$$

(where ψ is any of the functions $\delta H_{lp}^{(1)}$, $\delta H_{lp}^{(2)}$, and δH_p) cancel each other in the capillaries and cladding. Therefore, the result obtained is an analogue of the Ewald–Oseen extinction theorem [16]. Its physical validity is supported by considering the situations where system (32) admits analytical solutions. These situations satisfy the conditions

$$\varepsilon_l^{(1)} \neq \varepsilon_l^{(2)} = \varepsilon_s, \quad \varepsilon_k^{(1)} = \varepsilon_k^{(2)} = \varepsilon_s = \varepsilon_c$$

$$(k = 1, ..., l-1, l+1, ..., n), \tag{33}$$

$$\varepsilon_l^{(1)} = \varepsilon_l^{(2)} = \varepsilon_s \neq \varepsilon_c \quad (l = 1, 2, ..., n), \tag{34}$$

$$n = 1, \quad x_1 = y_1 = 0, \quad \varepsilon_1^{(1)} \neq \varepsilon_1^{(2)}, \quad \varepsilon_s \neq \varepsilon_c, \tag{35}$$

under which the fiber has circular symmetry. Under conditions (33) and (34), which correspond to circular rods surrounded by a homogeneous medium, system (32) in view of identity [20]

$$J_\nu(z)H_{\nu-1}^{(2)}(z) - J_{\nu-1}(z)H_\nu^{(2)}(z) = 2(\pi i z)^{-1} \tag{36}$$

is reduced to two homogeneous algebraic equations (for $\bar{A}_{l\nu}^{(0)}$ and $\bar{A}_{l\nu-2}^{(1)}$ subject to (33) or for $\bar{C}_\nu^{(0)}$ and $\bar{C}_{\nu-2}^{(1)}$ subject to (34)). In the case of (35), we are dealing with a four-layer fiber for which system (32) (in view of (36) and the equality $F_\nu^{(1)} = \delta_{\nu 0}$, where $\delta_{\nu 0}$ is the Kronecker delta) is reduced to four homogeneous algebraic equations for the unknowns $\bar{A}_{l\nu}^{(0)}$, $\bar{A}_{l\nu-2}^{(1)}$, $\bar{C}_\nu^{(0)}$, and $\bar{C}_{\nu-2}^{(1)}$. In all the cases mentioned above, the condition $\det M = 0$ results in closed dispersion relations for β. One may show that these equations are consistent with the well-known equations [18, 21], which were obtained by using the standard procedure of equating the tangential components of the fields at the interfaces.

In real microstructured fibers, conditions (33)–(35) fail and (32) represents an infinite set of algebraic equations. For engineering purposes, we leave only terms with $|\nu| \leq m$ in series (5) and (7), assuming that $A_{l\nu}^{(j)} = 0$ and $C_\nu^{(j)} = 0$ at $\nu > m$. This is equivalent to solving (1) by the standard method of quadratures [22]. In this case, the dimension of the matrix M is $N \times N$, where $N = 2(n+1)(2m+1)$. The value of the complex propagation constant β that satisfies the equation $\det M = 0$ can be found by contour integration [23]. As follows from calculations, the rank $\tilde{M}$ of the matrix equals $N - 1$ if $\det M = 0$. Therefore, we can express all the components of the vector X through one of them and then find the functions $H_x(x, y)$ and $H_y(x, y)$ inside the capillaries and cladding from (5), (7), and (29). In the intercapillary spacing,

$$\sqrt{(x - x_l)^2 + (y - y_l)^2} > b_l \quad (l = 1, 2, ..., n),$$

Calculated values of the propagation constant vs. order of reduction m

m	$k_0^{-1}\,\mathrm{Re}\beta$	$k_0^{-1}\,\mathrm{Im}\beta \times 10^6$	$k_0^{-1}\,\mathrm{Im}\beta_\mathrm{c} \times 10^6$	$k_0^{-1}\,\mathrm{Re}\beta_\mathrm{a}$	$k_0^{-1}\,\mathrm{Im}\beta_\mathrm{a} \times 10^6$
3	1.438470933	−1.439	−1.501	1.438528862	−6.918
5	1.438363682	−1.396	−1.381	1.438366726	−1.374
7	1.438364928	−1.417	−1.417	1.438364935	−1.416
9	1.438364934	−1.416	−1.416	1.438364934	−1.416
11	1.438364934	−1.416	−1.416	−	−

$$\sqrt{x^2 + y^2} < A,$$

these functions can be calculated by formulas (8), which (by virtue of (3), (5), (7), (9), (15)–(20), and (29)) are transformed into the form

$$H_x + (-1)^p i H_y$$

$$= \sum_{v=-\infty}^{\infty} \left[(T_{1v}^{(p)}\overline{C}_v^{(p)} + T_{2v}^{(p)}\overline{C}_{v+\sigma}^{(q)}) J_v(\chi_s\rho)\exp(iv\varphi) \right.$$

$$\left. + \sum_{l=1}^{n} (S_{lv}^{1p}\overline{A}_{lv}^{(p)} + S_{lv}^{2p}\overline{A}_{lv+\sigma}^{(q)}) H_v^{(2)}(\chi_s\rho_l)\exp(iv\varphi_l) \right]. \quad (37)$$

After the functions $H_x(x, y)$ and $H_y(x, y)$ have been determined, the final calculation of the vector field of the mode is performed by the formulas

$$H_z = (i\beta)^{-1}(\nabla_x H_x + \nabla_y H_y), \quad \mathbf{E} = (i\omega\varepsilon)^{-1}\nabla\times\mathbf{H}, \quad (38)$$

which follow from the Maxwell equations.

Note that the cross section of a microstructured fiber often contains orthogonal axes of symmetry. When these axes are aligned with the $0x$ and $0y$ coordinate axes, the field components of the mode satisfy the relationships [17]

$$H_x(-x, y) = \gamma_x H_x(x, y), \quad H_x(x, -y) = \gamma_x H_x(x, y), \quad (39)$$

$$H_y(-x, y) = -\gamma_y H_y(x, y), \quad H_y(x, -y) = -\gamma_y H_y(x, y), \quad (40)$$

where $\gamma_x^2 = \gamma_y^2 = 1$.

Qwing to the properties of (39) and (40) are even, the only independent unknowns in system (32) are the coefficients $\overline{C}_v^{(p)}$ and $\overline{A}_{lv}^{(p)}$, which refer to capillaries placed in the first quadrant. In addition, according to (5), (7), (39), and (40), we have

$$\overline{A}_{lv}^{(0)} = \gamma_x\overline{A}_{l-v}^{(1)} \quad \text{at} \quad x_l = 0, \quad y_l \geq 0,$$

$$\overline{A}_{lv}^{(0)} = (-1)^v\gamma_y\overline{A}_{l-v}^{(1)} \quad \text{at} \quad x_l \geq 0, \quad y_l = 0,$$

$$\overline{C}_v^{(0)} = \gamma_x\overline{C}_{-v}^{(1)} = (-1)^v\gamma_y\overline{C}_{-v}^{(1)},$$

$$\overline{C}_\mu^{(1)} = \overline{C}_\mu^{(0)} = 0, \quad \overline{A}_{l\mu}^{(1)} = \overline{A}_{l\mu}^{(0)} = 0 \quad \text{at} \quad x_l = y_l = 0,$$

where $\mu = \pm1, \pm3, \dots$ if $\gamma_x\gamma_y = 1$ or $\mu = 0, \pm2, \dots$ if $\gamma_x\gamma_y = -1$.

The features described above allow the dimension of the matrix M to be reduced to $N_1 \times N_1$. Here,

$$N_1 = (2n_1 + n_2)(2m + 1)$$

$$+ (n_3 + 1)\{m + 0.5[(-1)^m\gamma_x\gamma_y + 1]\},$$

where n_1 is the number of capillaries with $x_l > 0$ and $y_l > 0$; n_2 is the number of capillaries with $x_l = 0$, $y_l > 0$ and $x_l > 0$, $y_l = 0$; and n_3 is the number of capillaries with $x_l = y_l = 0$ (n_3 may take the value 0 or 1).

It is easy to check that $NN_1^{-1} \cong 4$. Thus, taking into account the symmetry of a microstructured fiber makes it possible to reduce significantly the body of computation, as demonstrated below.

To verify the method elaborated in this work, we evaluated a microstructured fiber with circular inclusions, which was studied in detail [14, 15] by the multipole method. The fiber has six air channels in quartz glass, which make up a hexagonal ring, and is a specific case of a capillary fiber if

$$\varepsilon_1^{(1)} = \varepsilon_1^{(2)} = \varepsilon_2^{(2)} = \dots = \varepsilon_7^{(2)} = \varepsilon_s,$$

$$\rho^{(1)} = 0, \quad \rho^{(2)} = \rho^{(3)} = \dots = \rho^{(7)} = \Lambda,$$

$$\varphi^{(2)} = 0, \quad \varphi^{(l+1)} - \varphi^{(l)} = \frac{\pi}{3} \quad (l = 2, 3, \dots, 6).$$

The table lists the values of the propagation constant of the fourth-order leaky mode (the fiber under consideration supports only leaky modes [14, 15]) that were obtained for $A = 14.25\ \mu m$, $\Lambda = 6.75\ \mu m$, $\lambda_0 = 1.45\ \mu m$, $a_l = 2.5\ \mu m$, $\varepsilon_l^{(1)} = 1$ ($l = 2, 3, \dots, 7$), $\varepsilon_s = 2.1025$, and $\varepsilon_c = 2.1025 - i2.9 \times 10^{-8}$ [16]. These values of β are the zeros of the determinant of the matrix M corresponding to $\gamma_x = 1$ and $\gamma_y = -1$. As follows from the table, the method of integral equations provides rapid convergence for the propagation constant. Similar convergence is observed for the mode field, as indicated by the values of $\mathrm{Im}\beta_c$, which are the imaginary part of the
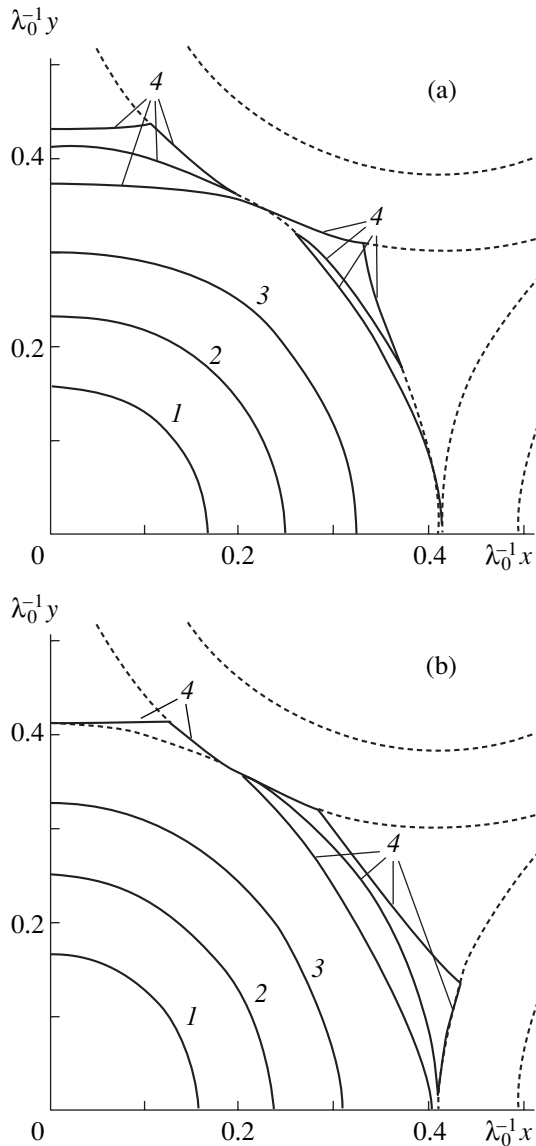
Fig. 2. Intensity isolines $S_{z\max}^{-1} S_z = (1)$ 0.8, (2) 0.6, (3) 0.4, and (4) 0.2 for the (a) H_x and (b) H_y modes of the capillary fiber. The dashed lines are the boundaries of the capillaries.

propagation constant calculated by the formula [17]

$$\mathrm{Im}\beta = -\left(2A^{-1}\int_0^A d\rho\rho\int_0^{2\pi} S_z d\varphi\right)^{-1} \int_0^{2\pi}(S_\rho)_{\rho = A}d\varphi,$$

where S_z and S_ρ are the longitudinal and radial components of the Poynting vector $\mathbf{S} = 0.5\mathrm{Re}(\mathbf{E} \times \mathbf{H}^*)$.

The values of β_a are the values of the propagation constant that were obtained by the multipole method [15]. They correspond to the order of reduction m in the expansions for the field longitudinal components E_z and H_z (expansions (5) for $\rho_l < a_l$) [14]. When m is sufficiently large, the propagation constants obtained by the method of integral equations and the multipole method

nearly coincide. The adequacy of the results obtained by the two independent approaches corroborates the validity of both of them.

DESIGN OF CAPILLARY FIBERS

Microstructured optical fibers are usually produced by drawing a macroscopic preform that has a set of capillaries [24]. Fibers thus produced always have intercapillary spacings of near-triangular shape. The design of the fiber depends greatly on the drawing conditions [24]. Two limiting designs (intercapillary spacings collapse due to the surface tension force or keep their original shape) are of greatest practical interest [24]. In the former case, well-studied [1, 3, 5, 8–15] fibers with circular inclusions arise. In the latter case, which readily occurs when the preform has thin-walled capillaries [4, 12, 24], capillary fibers are produced. Modes in capillary fibers have not yet been studied. The method of integral equations makes it possible to fill this gap. In this work, we applied this method in an attempt to reach a maximal contraction of the fundamental mode field distribution. The corresponding structures are of interest from the standpoint of observing nonlinear effects and bending loss minimization [4, 5].

Various capillary configurations (capillaries with different inner and outer radii embedded in glass tubes) were suggested in [5]. We will restrict our analysis to the hexagonal close packing of capillaries occupying the circular cross section of the cladding (Fig. 1b). In such fibers, the radii of the central rod and its nearest neighbors (capillaries of the inner ring) are the same: $b_1 = b_2 = \ldots = b_7$. The radii of the cladding and capillaries of the outer ring are related as

$$A = 3Kb_1, \quad b_8 = b_9 = \ldots = b_{13} = Kb_1,$$

$$K = \frac{1}{3}(1 + 2\sqrt{3} + 2\sqrt{1 + \sqrt{3}}).$$

Let us assume that (i) the central rod, capillaries, and cladding are made of the same material, quartz glass ($\varepsilon_1^{(1)} = \varepsilon_1^{(2)} = \varepsilon_2^{(2)} = \ldots = \varepsilon_{13}^{(2)} = \varepsilon_c$); (ii) intra- and intercapillary holes are filled with air ($\varepsilon_2^{(1)} = \varepsilon_3^{(1)} = \ldots = \varepsilon_{13}^{(1)} = \varepsilon_s = 1$); and (iii) the inner-to-outer radius ratios for the capillaries are the same ($a_2 b_2^{-1} = a_3 b_3^{-1} = \ldots = a_{13} b_{13}^{-1} = w$). Since the fields of the fiber modes penetrate into the air channels, the inequality $k_0^{-2} \mathrm{Re}\beta^2 < \mathrm{Re}\varepsilon_c$, which means that these modes are leaky, is met [8, 17]. Therefore, the outer capillary ring minimizes the fundamental mode leakage.

Figures 2 and 3 show results computed for the fields of the fundamental H_x and H_y modes of the fibers (the modes are designated by the basic component of the magnetic field) at $\varepsilon_c = 2.1025$. The plots were constructed based on expressions (5), (7), (29), (37), and
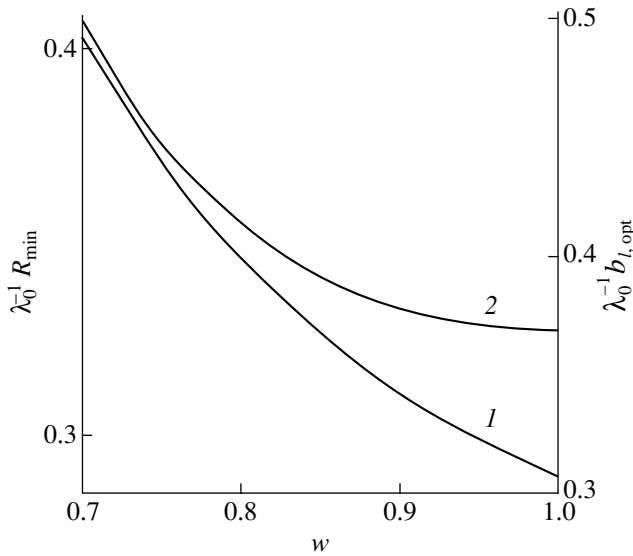
Fig. 3. (*1*) R_{min} and (*2*) $b_{l,\,opt}$ vs. w.

Fig. 4. (*1, 2*) Imβ, (*3*) D, and (*4*) R vs. λ_0 for the (*1, 3, 4*) fundamental and (*2*) first mode of the capillary fiber.

(38) after solving system (32) with $\gamma_x = \gamma_y = 1$ for the H_x mode and $\gamma_x = \gamma_y = -1$ for the H_y mode.

Figure 2 demonstrates the quarters the mode intensity symmetric distributions for the fiber with $w = 0.8$ and $b_1 = 0.4147\lambda_0$ that were obtained for $m \geq 12$ and correspond to $k_0^{-1}\beta = 1.300520 - i4.44 \times 10^{-10}$. The distributions are localized mainly within the inner capillary ring, the ratios $S_{z\,max}^{-1}S_z$ ($S_{z\,max} = S_z|_{\rho=0}$) becoming negligible at the boundary of the cladding ($S_{z\,max}^{-1}S_z|_{\rho=A} <$ 4.3×10^{-9} for the H_x mode and $S_{z\,max}^{-1}S_z|_{\rho=A} < 1.5 \times 10^{-9}$ for the H_y mode). Therefore, the transverse dimensions can be estimated in terms of the rms radius [15]

$$R = \sqrt{\left(\int_0^{2\pi}d\varphi\int_0^A\rho S_z d\rho\right)^{-1}\int_0^{2\pi}d\varphi\int_0^A\rho^3 S_z d\rho}.$$

Note that, although the intensity distributions of the H_x and H_y modes differ significantly (Fig. 2), the associated values of R are identical within the calculation accuracy. Specifically, the dependences for the H_x and H_y modes shown in Fig. 3 are indistinguishable on the scale of the figure, while for the H_x and H_y modes of the fiber to which Fig. 2 corresponds, we have $R = 0.345477\lambda_0$. In addition, the H_x and H_y modes of the capillary fibers under study are degenerate [14, 15], as follows from the theory of groups. This strict statement is consistent with our calculations, according to which the values of β for these modes coincided to 12 decimal places for $m \leq 25$. These properties allow us to describe the integral characteristics of the fundamental modes irrespective of their polarization.
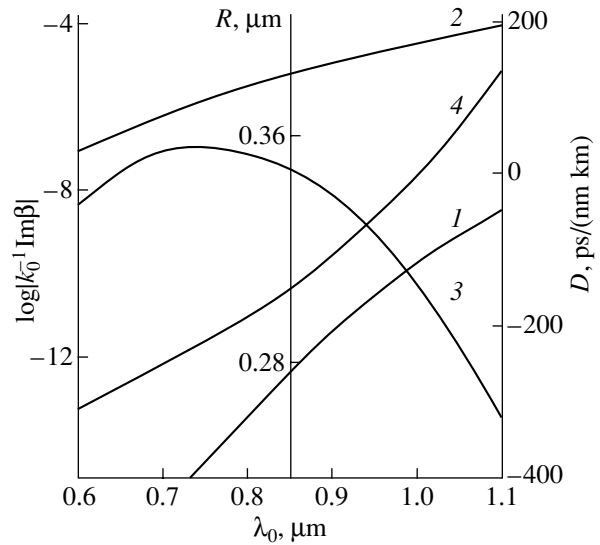
Figure 3 demonstrates the solution to the problem of finding minimal R for the fundamental mode. Here, R_{min} and $b_{1,\,opt}$ are the minimum value of R and the associated value of the argument of the function $R(b_1)$. As follows from Fig. 3, R_{min} and $b_{1,\,opt}$ are monotonically decreasing functions of w, and the minimum of the function $R_{min}(w)$ in the limit $w \longrightarrow 1$ (capillaries with infinitely thin walls) is physically unattainable. At a feasible value $w = 0.8$, the most contracted mode field distribution is observed in the fiber corresponding to Fig. 2.

The dispersion characteristics of the fiber with $w = 0.8$ and $b_1 = 0.429$ μm (this value coincides with $b_{1,\,opt}$ for $\lambda_0 = 1.0345$ μm) are shown in Fig. 4. They were calculated with allowance for the dispersion of quartz glass by using the Sellmeier trinomial formula [25]. In the approximation used, Im$\varepsilon_c = 0$; therefore, curves *1* and *2* in Fig. 4 characterize the mode attenuation due to leakage into the cladding.

As was noted, all modes in the fibers under study are leaky and the number of modes is infinitely large, since $\det M(\beta)$ is a transcendental function of finite type [26]. Hence, these fibers may be only quasi-single-mode fibers provided that leakage losses for the fundamental mode are much less than those for other modes. Of the higher modes, the first mode, which is an analogue of the TE_{01} mode of a homogeneous circular fiber [27] (for this mode, $\gamma_x = -1$ and $\gamma_y = 1$), is the lowest loss mode. Therefore, it will suffice to compare leakage losses for the fundamental and first modes. As follows from curves *1* and *2* in Fig. 4, the fiber may be considered quasi-single-mode throughout the range of λ_0. Specifically, at $\lambda_0 = 0.856$ μm, the attenuation of the fundamental and first modes is, respectively, 3.517×10^{-5} and 403.7 dB m^{-1}.

The dependence $D(\lambda_0)$ in Fig. 4 characterizes the dispersion of the group velocity v_g of the fundamental mode. Here,

$$D = \frac{dv_g^{-1}}{d\lambda_0} = -\frac{\lambda_0^2}{2\pi c}\frac{d^2\mathrm{Re}\beta}{d\lambda_0^2},$$

where c is the velocity of light in free space.

As follows from Fig. 4, the dispersion of the group velocity is zero at $\lambda_0 = 0.856$ μm. From curve *4* in Fig. 4, it follows that this λ_0 corresponds to $R = 0.308$ μm, which is close to $R_{min} = 0.296$ μm.

Thus, quasi-single-mode capillary fibers may combine the zero dispersion of the group velocity with a low attenuation and high spatial localization of the mode energy. Such a combination favors the observation of low-threshold nonlinear effects. Moreover, the mode spot area estimated from the value of R found is 0.3 μm^2. This area is 47 times smaller than the spot area for fibers with circular inclusions, which were used to generate a wide-band continuum in [4].

CONCLUSIONS

Our method of designing microstructured fibers, which is based on the analysis of integral equations, makes it possible to successively refine a solution to the vector waveguide problem with allowance for the leakage effect. This approach can be applied to optimize the mode characteristics of quartz capillary fibers close-packed in the circular cladding. Computation for a specific mode that is performed with a Pentium II (400 MHz) PC takes less than 50 s, which proves the efficiency of the method in designing microstructured fibers with desired properties. In this work, we only dealt with leaky modes, although normal eigenmodes (if supported by a microstructured fiber) are also easy to analyze. Associated results will be published later.

REFERENCES

1. T. A. Birks, J. C. Knight, and P. St. J. Russell, Opt. Lett. **22**, 961 (1997).
2. J. K. Ranka, R. S. Windeler, and A. J. Stentz, Opt. Lett. **25**, 25 (2000).
3. A. Ferrando and J. J. Miret, Appl. Phys. Lett. **78**, 3184 (2001).
4. T. M. Monro, W. Belardi, K. Furusawa, *et al.*, Meas. Sci. Technol. **12**, 854 (2001).
5. A. V. Belov and E. M. Dianov, Kvantovaya Élektron. (Moscow) **32**, 641 (2002).
6. C. Kerbage and B. J. Eggleton, Opt. Photonics News **13** (9), 38 (2002).
7. F. G. Omenetto, A. Efimov, A. J. Taylor, *et al.*, Opt. Express **11**, 61 (2003).
8. B. Kuhlmey, J. Renversez, and D. Maystre, Appl. Opt. **42**, 634 (2003).
9. T. M. Monro, D. J. Richardson, N. G. R. Broderick, *et al.*, J. Lightwave Technol. **17**, 1093 (1999).
10. D. Mogilevtsev, T. A. Birks, and P. St. J. Russell, J. Lightwave Technol. **17**, 2078 (1999).
11. A. Ferrando, E. Silvestre, J. J. Miret, *et al.*, J. Opt. Soc. Am. A **17**, 1333 (2000).
12. T. M. Monro, D. J. Richardson, N. G. R. Broderick, *et al.*, J. Lightwave Technol. **18**, 50 (2000).
13. Z. Zhu and T. G. Brown, Opt. Express **10**, 853 (2002).
14. T. P. White, B. T. Kuhlmey, R. C. McPhedran, *et al.*, J. Opt. Soc. Am. B **19**, 2322 (2002).
15. B. T. Kuhlmey, T. P. White, G. Renversez, *et al.*, J. Opt. Soc. Am. B **19**, 2331 (2002).
16. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1969; Nauka, Moscow, 1973).
17. A. B. Sotsky and L. I. Sotskaya, Opt. Spektrosk. **88**, 465 (2000) [Opt. Spectrosc. **88**, 415 (2000)].
18. D. Marcuse, *Light Transmission Optics* (Van Nostrand, New York, 1972; Mir, Moscow, 1974).
19. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1984).
20. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
21. P. Yeh, A. Yariv, and E. Marom, J. Opt. Soc. Am. **68**, 1196 (1978).
22. V. I. Krylov, V. V. Bobkov, and P. I. Monastyrskiĭ, *Computational Methods* (Nauka, Moscow, 1977), Vol. 2.
23. A. B. Sotsky, Dokl. Akad. Nauk Belarusi **45** (3), 19 (2001).
24. A. D. Fitt, K. Furusawa, T. M. Monro, *et al.*, J. Lightwave Technol. **19**, 1924 (2001).
25. M. Adams, *An Introduction to Optical Waveguides* (Wiley, New York, 1981; Mir, Moscow, 1984).
26. V. V. Shevchenko, Differentsial'nye Uravneniya **15**, 2004 (1979).
27. B. T. Kuhlmey, R. C. McPhedran, C. M. de Sterke, *et al.*, Opt. Express **10**, 1285 (2002).

Translated by V. Isaakyan

GASES
AND LIQUIDS

Relationship between the Diffusiophoresis Rate, Evaporation Coefficient, and Size of a Large Volatile Drop

G. Yu. Yalamov

Sholokhov State Open Pedagogical University, Institute of Informatization in Education,
Moscow, 109004 Russia
e-mail: aio.net@relcom.ru
Received June 17, 2003

Abstract—A new theory of diffusiophoresis of large volatile spherical aerosol drops that is an extension of investigations [1–8] is developed. The influence of the radius of the drop, the surface tension coefficient varying over the surface of the drop, the evaporation coefficient α of the liquid, and the flows inside the drop on the diffusiophoresis rate are taken into account. Expressions obtained allow for direct determination of the velocity of large individual aerosol drops in a binary gas mixture nonuniform in component concentration. It is shown that both the magnitude and the direction of the diffusiophoresis velocity depend on α and the size of the drop. It is assumed that the size of the drop varies but remains considerably greater than the mean free path of gas molecules. © 2004 MAIK "Nauka/Interperiodica".

Drops (particles) are called volatile if the constituent material evaporates or condenses on their surface [5–7]. Otherwise (in the absence of the surface phase transition), drops are nonvolatile.

If the radius R of a spherical aerosol particle far exceeds the mean free path $\bar{\lambda}$ of environmental molecules, the particle is called large. For such particles, the Knudsen number

$$\bar{\lambda}/R \ll 1. \tag{1}$$

For example, water drops suspended in air at normal conditions are large if their radius $R \geq 6$ μm.

Diffusiophoresis of large nonvolatile and volatile particles has been covered in [2–4] and [1, 5–8], respectively. In [1, 5–7], the direct influence of the evaporation coefficient α of the liquid on the diffusiophoresis rate was neglected. In [8], the effect of the evaporation coefficient α was considered without regard for internal flows. However, it was shown [4, 6] that internal flows make a significant contribution to the diffusiophoresis rate if the viscosity of the liquid inside the drop is comparable to that of the environment. Therefore, a need has arisen for a theory of diffusiophoresis of large spherical volatile aerosol particles that would directly take into account the effect of the evaporation coefficient α when flows are present inside the drop and the interface surface tension varies over its surface. The effect of the size of a large drop on the diffusiophoresis rate at a constant evaporation coefficient is also of interest.

Let a spherical drop of radius R contain a one-component liquid with a thermal conductivity χ_i and molecular mass m_1. The drop is immersed in a binary gas mixture nonuniform in component concentration. One of the components is the vapor of the drop liquid. The mixture has a thermal conductivity χ_o, viscosity η_{0o}, and interdiffusion coefficient $D_{12}^{(o)}$ (hereafter, the indices o and i refer to quantities outside and inside the drop, respectively). Far away from the drop, the gradients of the gas component concentrations $(\nabla C_{1o})_\infty$ and $(\nabla C_{2o})_\infty$ are kept constant. Here, C_{1o} and C_{2o} are the relative numerical concentrations defined as

$$C_{1o} = \frac{n_{1o}}{n_o} \quad \text{and} \quad C_{2o} = \frac{n_{2o}}{n_o}, \tag{2}$$

where n_{1o} and n_{2o} are the numbers of molecules of the gas mixture components per unit volume and $n_o = n_{1o} + n_{2o}$.

Evidently, $C_{1o} + C_{2o} = 1$ and the condition

$$\nabla C_{1o} = -\nabla C_{2o} \tag{3}$$

is met at any point of the gas mixture.

The radius of the drop is assumed to be large compared to the mean free paths λ_1 and λ_2 of molecules of either component so that the problem may be solved in the hydrodynamic approximation [1–7]. It is also assumed that the moving drop retains its spherical shape. According to [7], this holds true if its external pressure force is small compared to the surface tension force:

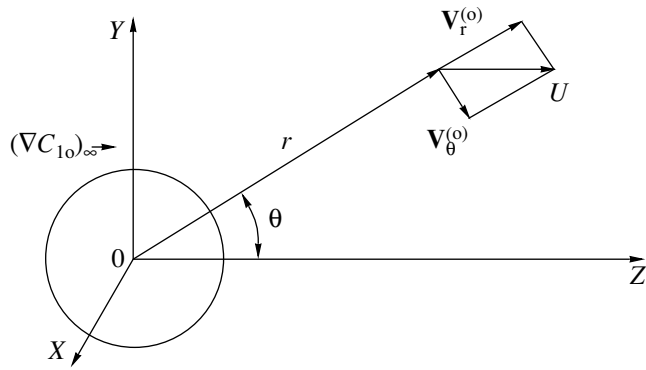$$\sigma/R \ll \eta_{0o}\frac{|\mathbf{U}|}{R}. \tag{4}$$

Fig. 1.

Here, σ is the surface tension coefficient at the drop–gas interface, R is the radius of the drop, and $|\mathbf{U}|$ is the absolute velocity of the gas mixture relative to the drop.

Since we assume that the spherical shape of the drop is preserved during the motion, it is convenient to solve the problem in the spherical coordinate system (r, θ, φ) with the origin at the center of the drop.

Let the polar axis $z = r\cos\theta$ be aligned with the vector $(\nabla C_{1o})_\infty$. The drop is at rest, and the center of gravity of the external gas mixture moves with respect to the center of the drop with a velocity $\mathbf{U}$ at $r \longrightarrow \infty$ (Fig. 1) [7, 8].

The distributions of velocities, pressures, temperatures, and concentrations outside and inside the drop satisfy the following system of differential linearized vector equations [7]:

$$\eta_{0o}\nabla^2\mathbf{v}^{(o)} = \nabla p^{(2)}, \tag{5}$$

$$\mathrm{div}\,\mathbf{v}^{(o)} = 0, \tag{6}$$

$$\eta_{0i}\nabla^2\mathbf{v}^{(i)} = \nabla p^{(i)}, \tag{7}$$

$$\mathrm{div}\,\mathbf{v}^{(i)} = 0, \tag{8}$$

$$\nabla^2 C_{1o} = 0, \tag{9}$$

$$\nabla^2 T_o = 0, \tag{10}$$

$$\nabla^2 T_i = 0. \tag{11}$$

In Eqs. (5)–(11), $\mathbf{v}^{(o)}$ and $\mathbf{v}^{(i)}$ are the velocities of the center of inertia of the mixture outside and inside the drop, respectively; $p^{(o)}$ and $p^{(i)}$ are the respective pressures; η_{0o} and η_0 are the mean viscosities of the gas mixture and the drop, respectively; C_{1o} is the relative concentration of the first component outside the drop; and T_o and T_i are the temperatures outside and inside the drop, respectively.

Far away from the drop ($r \longrightarrow \infty$), the boundary conditions have the form [7, 8] (see also Fig. 1)

$$v_r^{(o)} = |\mathbf{U}|\cos\theta, \tag{12}$$

$$v_\theta^{(o)} = -|\mathbf{U}|\sin\theta, \tag{13}$$

$$p^{(o)} = p_0^{(o)}, \tag{14}$$

$$C_{1o} = C_{01o} + |(\nabla C_{1o})_\infty|r\cos\theta, \tag{15}$$

$$T_o = T_{0o}. \tag{16}$$

The boundary conditions on the surface of the drop are

$$\left(n_{02o}v_r^{(o)} - D_{12}^{(o)}n_{0o}^2\frac{m_1}{\rho_{0o}}\frac{\partial C_{2o}}{\partial r}\right)\bigg|_{r=R} = 0, \tag{17}$$

$$\left(n_{01o}v_r^{(o)} - D_{12}^{(o)}n_{0o}^2\frac{m_1}{\rho_{0o}}\frac{\partial C_{1o}}{\partial r}\right)\bigg|_{r=R}$$
$$= n_{0o}\alpha v(C_{1o}^{(H)} - C_{1o})\big|_{r=R}, \tag{18}$$

$$v_\theta^{(o)} - v_\theta^{(i)}\big|_{r=R} = \frac{K_{TC}^{(o)}}{T_{0o}R}\frac{\partial T_o}{\partial\theta}\bigg|_{r=R} + \frac{K_{DC}D_{12}^{(o)}}{R}\frac{\partial C_{1o}}{\partial\theta}\bigg|_{r=R}, \tag{19}$$

$$T_o\big|_{r=R} = T_i\big|_{r=R}, \tag{20}$$

$$\left(-\chi_o\frac{\partial T_o}{\partial r} + \chi_i\frac{\partial T_i}{\partial r}\right)\bigg|_{r=R}$$
$$= -n_{0o}\alpha vLm_1(C_{1o}^{(H)} - C_{1o})\big|_{r=R}. \tag{21}$$

The temperature T_i inside the drop and the concentration C_{1o} must be finite.

Boundary condition (17) reflects the fact that the surface of the drop is impermeable to the second component of the mixture, which does not experience a phase transition. In (17), the first and the second terms correspond to the radial convective and diffusion flows of the second component, respectively. Equation (18) means the continuity of the radial flow of the first (volatile) component at the surface of the drop. Its left-hand side represents the total radial flow of this component outside the drop (the sum of the convective and diffusive flows). The right-hand side of (18) is the radial flow that is removed from the surface through the Knudsen layer and is proportional to the evaporation coefficient α of the liquid in the drop. Previously, the latter flow was disregarded in the theory of diffusiophoresis of large volatile aerosol particles [5–7]. The expression for this term,

$$n_{0o}\alpha v(C_{1o}^{(H)} - C_{1o})\big|_{r=R},$$

is derived from simple statistical considerations [9]. Here,

$$v = \sqrt{\frac{kT_{0o}}{2\pi m_1}}$$

is quarter the mean absolute thermal velocity of vapor molecules, n_{0o} is the mean number of mixture molecules per unit volume, $C_{1o}^{(H)}$ is the relative concentration of the volatile saturated vapor at the surface of the drop, and k is the Boltzmann constant.

In Eqs. (17) and (18), n_{01o} and n_{02o} are the mean numbers of mixture component molecules per unit volume, m_1 and m_2 are the molecular masses of the components, $\rho_{0o} = n_{01o}m_1 + n_{02o}m_2$ is the mean mass density of the mixture, and $n_{0o} = n_{01o} + n_{02o}$.

Condition (19) describes the well-known phenomena of thermal and diffusion creep of a binary gas mixture over the surface of the drop. The rates of these phenomena depend on the thermal, $K_{TC}^{(o)}$, and diffusion, $K_{DC}^{(o)}$, creep coefficients [7, 8, 10–16].

The temperature at the drop–gas interface is continuous and given by Eq. (20). The continuity condition for the heat flux at the interface is given by relationship (21). The right of Eq. (21) takes into account the heat of phase transition, which is proportional to the heat of evaporation (condensation) L of the first gas component.

The boundary conditions mentioned above must be complemented by the condition of zero radial component of the convective flow of the liquid across the surface of the drop,

$$v_r^{(i)} = 0 \quad \text{at} \quad r = R, \tag{22}$$

and the continuity conditions for the radial and tangential components of the viscous stress tensor on the surface of the drop [7]:

$$\left(-p^{(o)} + 2\eta_{0o}\frac{\partial v_r^{(o)}}{\partial r}\right)\Bigg|_{r=R} - \frac{2\sigma_0}{R} - 2\frac{\partial \sigma}{\partial T_i}\Bigg|_{T_i - T_{0i}}$$
$$\times (T_i - T_{0i})|_{r=R} = \left(p^{(i)} + 2\eta_{0i}\frac{\partial v_r^{(i)}}{\partial r}\right)\Bigg|_{r=R}, \tag{23}$$

$$\eta_{0o}\left(\frac{1}{r}\frac{\partial v_r^{(o)}}{\partial \theta} + \frac{\partial v_\theta^{(o)}}{\partial r} - \frac{v_\theta^{(o)}}{r}\right)\Bigg|_{r=R} + \frac{1}{r}\frac{\partial \sigma}{\partial T_i}\Bigg|_{T_i = T_{0i}}\frac{\partial T_i}{\partial \theta}\Bigg|_{r=R}$$
$$= \eta_{0o}\left(\frac{1}{r}\frac{\partial v_r^{(i)}}{\partial \theta} + \frac{\partial v_\theta^{(i)}}{\partial r} - \frac{v_\theta^{(i)}}{r}\right)\Bigg|_{r=R}. \tag{24}$$

In [24], σ_0 is the mean value of the surface tension coefficient at the drop–gas interface. Representing the surface tension σ in Eqs. (23) and (24) in terms of the

small parameter $|(R\nabla C_{1o})_\infty|$ and leaving only the linear term, we obtain

$$\sigma = \sigma_0 + \frac{\partial \sigma}{\partial T_i}\Bigg|_{T_i = T_{0i}}(T_i - T_{0i}). \tag{25}$$

Note that the surface tension decreases with temperature; i.e., $\partial\sigma/\partial T < 0$ as follows from experimental data [17] and the simple theory concerned with this issue [18, 19].

It should be emphasized that boundary conditions (17)–(24) are linearized in the small parameter mentioned above and the quantities n_o, n_{1o}, n_{2o}, η_o, η_i, and T_o are replaced by their mean values [7].

Note also that the saturated concentration $C_{1o}^{(H)}$ is a function of the temperature T_i and can also be expanded in small parameter $|(R\nabla C_{1o})|$. Leaving only the linear terms yields

$$C_{1o}^{(H)}(T_i)|_{r=R} = C_{01o}^{(H)}(T_{0i})|_{r=R}$$
$$+ \frac{\partial C_{1o}^{(H)}}{\partial T_i}\Bigg|_{T_i = T_{0i}}(T_i - T_{0i})|_{r=R}. \tag{26}$$

The problem of the gas mixture flowing about the drop and internal flows have azimuthal symmetry, since the polar axis is aligned with the vector $(\nabla C_{1o})_\infty$. Therefore, the variables $\mathbf{v}^o$, $\mathbf{v}^{(i)}$, $p^{(o)}$, $p^{(i)}$, C_{1o}, T_o, and T_i are independent of the azimuth angle φ and $v_\varphi^{(o)}$ and $v_\varphi^{(i)}$ vanish [3–7].

In the spherical coordinates, a solution to set (5)–(11) with boundary conditions (12)–(15) can be represented in the form [7]

$$\mathbf{v}_r^{(o)} = \left(\frac{A_o}{r^3} + \frac{B_o}{r} + |\mathbf{U}|\right)\cos\theta + \frac{\gamma_o}{r}, \tag{27}$$

$$\mathbf{v}_\theta^{(o)} = \left(\frac{A_o}{2r^3} - \frac{B_o}{2r} - |\mathbf{U}|\right)\sin\theta, \tag{28}$$

$$p^{(o)} = p_0^{(o)} + \eta_{0o}\frac{B_o}{r^2}\cos\theta, \tag{29}$$

$$C_{1o} = C_{01o} + |(\nabla C_{1o})|_\infty r\cos\theta + \frac{\mu_o^{(c)}}{r^2}\cos\theta + \frac{\varphi^{(C)}}{r}, \tag{30}$$

$$T_o = T_{0o} + \frac{\mu_o^{(T)}}{r^2}\cos\theta + \frac{\varphi^{(T)}}{r}, \tag{31}$$

$$\mathbf{v}_r^{(o)} = (Q_i + D_i r^2)\cos\theta + \mathbf{v}_{r0}^{(i)}, \tag{32}$$

$$\mathbf{v}_\theta^{(i)} = -(Q_i + 2D_i r^2)\sin\theta, \tag{33}$$

$$p^{(i)} = p_0^{(i)} + 10\eta_i D_i r^2\sin\theta, \tag{34}$$

$$T_i = T_{0i} + \mu_i^{(T)}r\cos\theta. \tag{35}$$

Next, substituting solutions (27)–(35) into boundary conditions (17)–(24) yields a system of algebraic equations for the unknown constants A_o, B_o, $|\mathbf{U}|$, $\mu_o^{(C)}$, $\mu_o^{(T)}$, $\mu_i^{(T)}$, Q_i, D_i, $\varphi^{(c)}$, φ^T, $v_{r0}^{(i)}$, and C_{01o}.

To proceed further, we need an analytical expression for $|\mathbf{U}|$, where $\mathbf{U}$ is the velocity of the incoming gas mixture. Then, the diffusiophoresis velocity vector $\mathbf{U}_D$ with respect to the center of gravity of the gas mixture [7] can be written as

$$\mathbf{U}_D = -\mathbf{U}. \tag{36}$$

Using the expression for $\mathbf{U}$ (in vector form), we obtain

$$\mathbf{U}_D$$

$$-\frac{6\eta_{0i}}{(3\eta_{0i}+2\eta_{0o})\Delta}\left\{K_{DC}^{(o)}D_{12}^{(o)2}[2\chi_o + \chi_i + n_{0o}\alpha v L m_1 R\delta]\right.$$
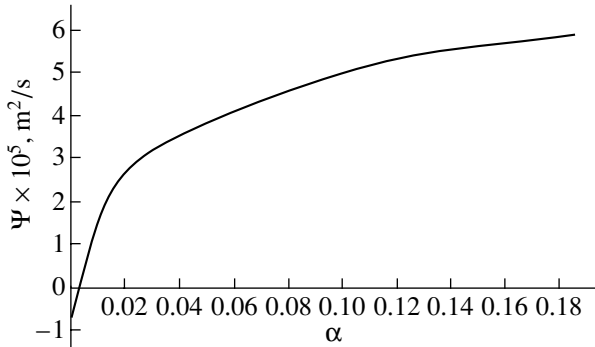
Fig. 2. Proportionality coefficient Ψ vs. the coefficient of evaporation α under normal conditions.
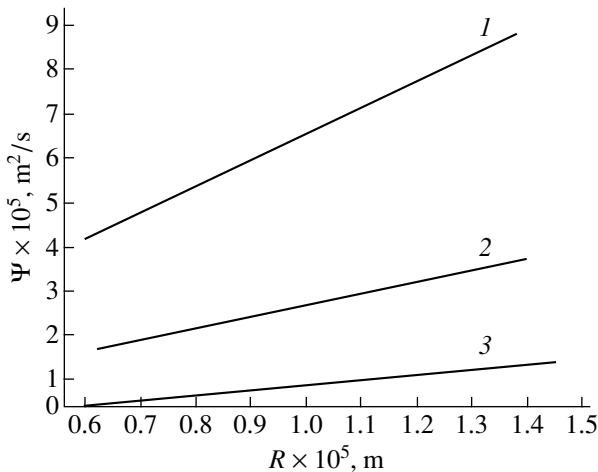
Fig. 3. Coefficient Ψ vs. the radius R of the drop at normal conditions.

$$+\left[\frac{K_{TC}^{(o)}}{T_{0o}} + \frac{R}{3\eta_{0i}}\delta_\sigma\right]D_{12}^{(o)}n_{0o}\alpha v L m_1 R\right\}(\nabla C_{1o})_\infty \tag{37}$$

$$+\frac{3(\eta_{0i}+2\eta_{0o})D_{12}^{(o)}m_1}{(3\eta_{0i}+2\eta_{0o})\Delta\rho_{0o}}\{[2\chi_o + \chi_i]n_{0o}\alpha v R\}(\nabla C_{1o})_\infty,$$

where

$$\Delta = [2\chi_o + \chi_i]$$
$$\times\left(2D_{12}^{(o)} + \frac{n_{02o}}{n_{0o}}\alpha v R\right) + 2D_{12}^{(o)}n_{0o}\alpha v L m_1 R\delta, \tag{38}$$

$$\delta = \left.\frac{\partial C_{1o}^{(H)}}{\partial T_i}\right|_{T_i = T_{0i}}, \quad \delta_\sigma = \left.\frac{\partial\sigma}{\partial T_i}\right|_{T_i = T_{0i}}. \tag{39}$$

Consider a number of limiting cases for Eq. (37). When $\alpha \longrightarrow 0$, i.e., when the coefficient of evaporation of the liquid in the drop tends to zero, we have

$$\lim_{\alpha\to 0}\mathbf{U}_D = -\frac{3\eta_{0i}}{3\eta_{0i}+\eta_{0o}}K_{DC}^{(o)}D_{12}^{(o)}(\nabla C_{1o})_\infty. \tag{40}$$

This formula gives the diffusiophoresis velocity for a large drop with regard for internal flows. If $\eta_{0i} \gg \eta_{0o}$, which is quite feasible, we derive from (40) at $\eta_{0o}/\eta_{0i} \longrightarrow 0$

$$\lim_{\substack{\alpha\to 0 \\ \eta_{0o}/\eta_{0i}\to 0}}\mathbf{U}_D = -K_{DC}^{(o)}D_{12}^{(o)}(\nabla C_{1o})_\infty, \tag{41}$$

which is the classical expression for the diffusiophoresis velocity of a large nonvolatile drop (particle).

If the thermal conductivity of the drop $\chi_i \gg \chi_o$, i.e., in the limit $\chi_o/\chi_i \longrightarrow 0$, Eq. (37) yields

$$\lim_{\chi_o/\chi_i\to 0}\mathbf{U}_D = -\frac{6\eta_{0i}}{(3\eta_{0i}+2\eta_{0o})\Delta_1}$$

$$\times\left\{K_{DC}^{(o)}D_{12}^{(o)2}\left[1 + \frac{n_{0o}\alpha v L m_1 R\delta}{\chi_i}\right]\right.$$

$$+\left[\frac{K_{TC}^{(o)}}{T_{0o}} + \frac{R}{3\eta_{0i}}\delta_\sigma\right]\frac{D_{12}^{(o)}n_{0o}\alpha v L m_1 R}{\chi_i}\right\}(\nabla C_{1o})_\infty \tag{42}$$

$$+\left\{\left(\frac{3(\eta_{0i}+2\eta_{0o})}{3\eta_{0i}+2\eta_{0o}}\right)\frac{D_{12}^{(o)}m_1}{\rho_{0o}\Delta_1}n_{0o}\alpha v R\right\}(\nabla C_{1o})_\infty,$$

where

$$\Delta_1 = \left(2D_{12}^{(o)} + \frac{n_{02o}}{n_{0o}}\alpha v R\right) + \frac{2D_{12}^{(o)}n_{0o}\alpha v L m_1 R\delta}{\chi_i}.$$

The estimates made with the formulas derived above indicate that an increase in the coefficient of evaporation α of the liquid changes both the magnitude and the direction of the diffusiophoresis velocity $\mathbf{U}_D$.

This is clearly seen in Fig. 2, where the proportionality factor Ψ between $\mathbf{U}_D$ and $(\nabla C_{1o})_\infty$ ($\mathbf{U}_D = \Psi(\nabla C_{1o})_\infty$) is plotted versus the coefficient of evaporation of a water drop with a radius of 10 μm suspended in air under normal conditions. If α is very small ($0 < \alpha < 0.003$), the velocity $\mathbf{U}_D$ decreases, retaining the initial direction ($\mathbf{U}_D < 0$). This is because the effects of thermal and diffusion creep (due to the nonuniform phase transition over the drop surface) dominate in this case and the drop moves towards lower concentrations C_{1o} when $K_{DC} > 0$. When $\alpha \geq 0.003$, the effect of purely reactive evaporation and variable surface tension ($\partial\sigma/\partial T_i < 0$) starts prevailing and the drop moves towards higher C_{1o}.

On the other hand, at constant α, large drops of various diameter have different velocities $\mathbf{U}_D$, as demonstrated by curves *1–3* in Fig. 3. These curves show Ψ vs. the radius R of a water drop suspended in air under normal conditions for $\alpha = 0.5$ (curve *1*), 0.1 (curve *2*), and 0.008 (curve *3*). With very small values of α (<0.01), drops with a certain radius R may have the diffusiophoresis velocity $\mathbf{U}_D = 0$. For example, curve *3* shows that the drops with a radius of 6 μm are motionless.

Thus, knowing the radius of a motionless drop under given external conditions (T_{0o}, $P_0^{(o)}$), one can determine the coefficient of evaporation of the liquid in the drop, which is of interest for experimental observation.

REFERENCES

1. B. V. Derjaguin and Yu. I. Yalamov, in *International Reviews on Aerosol Physics and Chemistry* (Pergamon, Oxford, 1972), Vol. 3, Part 2, pp. 1–200.

2. I. R. Brock, J. Colloid Sci. **18**, 489 (1963).

3. Yu. I. Yalamov and B. A. Obukhov, Zh. Tekh. Fiz. **42**, 1064 (1979) [Sov. Phys. Tech. Phys. **42**, 844 (1979)].

4. Yu. I. Yalamov and A. S. Sanasaryan, Zh. Tekh. Fiz. **47**, 1063 (1977) [Sov. Phys. Tech. Phys. **22**, 634 (1977)].

5. Yu. M. Agvanyan and Yu. I. Yalamov, Kolloidn. Zh. **40**, 1043 (1978).

6. Yu. M. Agvanyan and Yu. I. Yalamov, Inzh.-Fiz. Zh. **37**, 1083 (1979).

7. Yu. I. Yalamov and V. S. Galoyan, *Dynamics of Drops in Inhomogeneous Viscous Media* (Luĭs, Yerevan, 1985).

8. G. Yu. Yalamov, in *Topical Problems of Physics and Mechanics of Aerodisperse Systems: Collection of Articles from the Krupskaya Moscow Regional Pedagogical Institute* (MOPI, Moscow, 1989), pp. 128–140.

9. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 1.

10. I. N. Ivchenko and Yu. I. Yalamov, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 6, 59 (1969).

11. H. A. Kramers and I. Kistemaker, Physica **10**, 699 (1943).

12. R. Ya. Kucherov, Zh. Tekh. Fiz. **27**, 2158 (1957) [Sov. Phys. Tech. Phys. **2**, 2001 (1957)].

13. Yu. I. Yalamov, I. N. Ivchenko, and B. V. Deryagin, Dokl. Akad. Nauk SSSR **180**, 330 (1968).

14. I. P. Breton, Phys. Fluids **12**, 2019 (1969).

15. Yu. Yu. Abramov and G. G. Gladush, Prikl. Mekh. Tekh. Fiz., No. 4, 51 (1970).

16. I. N. Ivchenko and Yu. I. Yalamov, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 4, 22 (1971).

17. N. B. Vargaftik, *Tables of the Thermophysical Properties of Liquids and Gases* (Nauka, Moscow, 1972; Halsted, New York, 1975).

18. E. A. Shtrauf, *Course of Physics* (Leningrad, 1961), Vol. 1, Part 3, Chap. 3.

19. R. V. Telesnin, *Molecular Physics* (Vysshaya Shkola, Moscow, 1973), Chap. 10.

Translated by A. Sidorova

GASES
AND LIQUIDS

Crisis of Hydrodynamic Drag of Drops in the Two-Phase Turbulent Flow of a Spray Produced by a Mechanical Nozzle at Transition Reynolds Numbers

N. N. Simakov

Yaroslavl State Technical University, Yaroslavl, 150023 Russia
e-mail: simakov@ystu.yar.ru
Received April 11, 2003; in final form, June 2, 2003

Abstract—When processing experimental data for the hydrodynamics of a two-phase flow in a spray produced by a mechanical nozzle, we revealed an anomaly in the behavior of the hydrodynamic drag of drops: the drag coefficient turns out to be four to seven times lower than the previously known values. Several hypotheses are put forward to explain the anomaly. It is found that, when the gas flows around drops under highly turbulent conditions, an "early" (i.e., observed even at transition Reynolds numbers, Re > 50) crisis of drag resistance of drops takes place. This new physical phenomenon allows us to account for a number of features of the two-phase flow that are observed in the experiment. Among these features is, in particular, the fact that the momentum transferred to the gas is roughly half the initial momentum of the liquid jet. © 2004 MAIK "Nauka/Interperiodica".

The effect of a drastic three- or fourfold decrease in the hydrodynamic drag C_d of a sphere, cylinder, or any other high-drag body, which occurs at Re on the order of 10^5 and is called the crisis of drag resistance, has been known for a long time [1–3]. It arises when the laminar boundary layer separates from the body surface, becomes turbulent, and the line of separation shifts downstream of the flow toward the "stern" region. When this takes place, the pressure profile in the flow changes and the flow about the body approaches a perfectly streamlined flow pattern [1, 2].

In the experimental study [4] of the hydrodynamics of a two-phase flow in a spray produced by a centrifugal jet nozzle (designed by the All-Russia Institute of Heat Engineering) with an outlet diameter of 2 mm, the crisis of drag reduction was observed for transition Reynolds numbers, Re = 40–130.

In our experiments, water was atomized in air by a nozzle placed vertically downward. We measured the dispersity of the spray (the drop size spectrum and mean Sauter diameter d_{32}), the radial velocity profile, the concentrations and specific flows of the liquid, the velocities and pressures of the gas at different distances (up to 1 m) from the nozzle and various overpressures (p = 300, 500, and 900 kPa) before the nozzle, and the variation of these parameters along the spray axis (the measurement error was within 5%).

The dispersity of the spray was measured by the method of small-angle light scattering [5]. At p = 500 kPa, the mean volume-to-surface ratio (the diameter d_{32}) of the drops was found to be equal to about 140 μm and varied in inverse proportion to the square root of the pressure at the nozzle. Its decrease with height of the spray was insignificant (12%).

The velocities of the drops were measured using the one-beam time-of-flight laser method [6] in its improved version [4]; the pressure and velocity of the gas were measured with the pneumatic-type metering technique using modified Pitot–Prandtl tubes and a high-sensitive capillary microgauge [4]. It turned out that, at each of the measurement points, the velocities of the drops are distributed in a wide range: their variance amounts to 25% of the mean value.

Figures 1–4 show, respectively, the radial profiles of the axial component $U_z(r, z)$ of the mean velocity of the drops, the gas velocity field $\mathbf{W}(r, z)$ at p = 500 kPa, and the variation of these parameters with the height of the spray (at $r = 0$) for p = 300, 500, and 900 kPa.

Figure 5 demonstrates the curves approximating the experimental dependences $U_z(0, z)$ and $W(0, z)$ at p = 500 kPa, as well as the curves for the relative velocity $W_{rel} = U - W$ of the phases and for the acceleration a_z of the drops at the spray axis. The acceleration is calculated by the formula

$$a_z = \frac{dU_z}{dt} = \frac{dU_z}{dz}\frac{dz}{dt} = U_z\frac{dU_z}{dz} = \frac{1}{2}\frac{dU_z^2}{dz}. \quad (1)$$

Using the data in Fig. 5, knowing the diameter $d = d_{32}$ of the drop and the hydrodynamic drag force $F = ma_z$, and neglecting the force of gravity mg, one can determine the drag coefficient C_d from the well-known
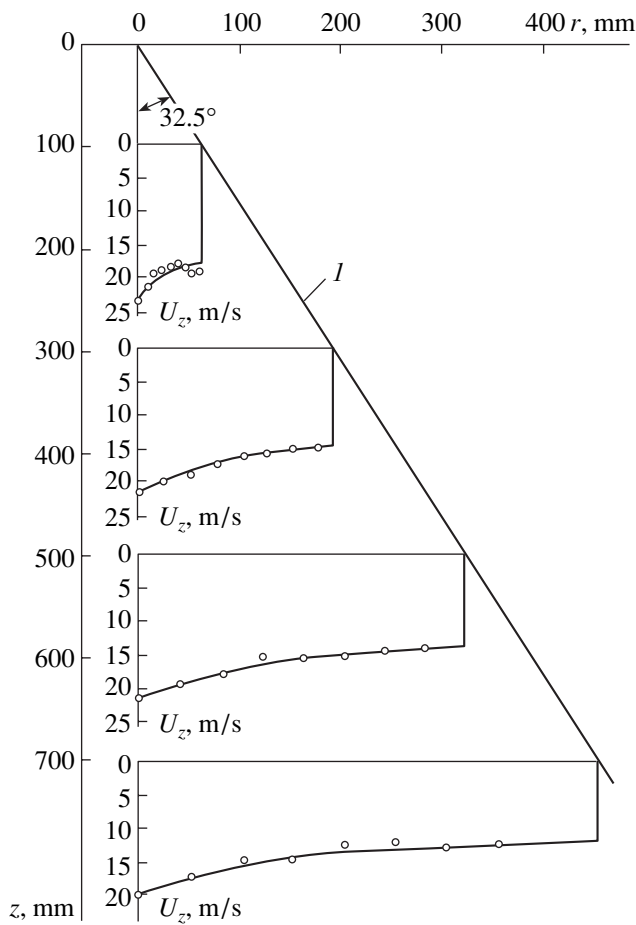
Fig. 1. Radial profiles of the axial velocity of the drops for the pressure $p = 50$ kPa at the nozzle. (*1*) Is the spray boundary.
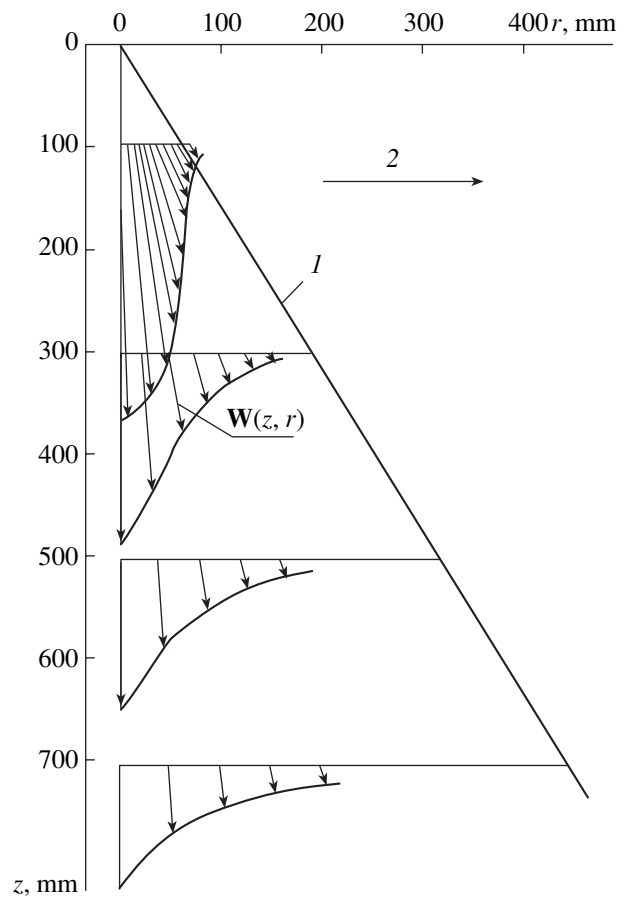

Fig. 2. Gas velocity field for the pressure $p = 500$ kPa at the nozzle. (*1*) The spray boundary and (*2*) velocity vector scale $|\mathbf{W}| = 10$ m/s.

formula

$$F = C_{d}S\rho(W_{rel})^{2}/2, \tag{2}$$

where ρ is the gas (air) density and $S = \pi d^2/4$ is the midsectional area of the drops.

For the Re range used in our experiment with a spray from the nozzle (Re = 40–130), the drag coefficient of a sphere may be approximated in different ways, e.g., by the widely used Klyachko formula [7–9] for solitary spherical particles

$$C_{d} = 24/\text{Re} + 4/\text{Re}^{1/3}. \tag{3}$$

Empirical formulas are also known for the volumetric density of the total force F_{Σ} of drop–gas interfacial interaction, such as the Ergun formula [10], which was used in [11, 12]. For low-density disperse flows, the modification of this formula [13, 14] is more appropriate:

$$dF_{\Sigma}/dV = \alpha(18\mu/(d_{32})^{2} + 0.36\rho/d_{32}W_{rel})W_{rel}, \tag{4}$$

where V is the volume of the liquid, α is the volumetric fraction of the liquid, and μ is the viscosity coefficient of the gas. For the drag coefficient of an individual particle, the modified formula gives

$$C_{d} = 24/\text{Re} + 0.48. \tag{5}$$

Figures 6 and 7 plot the drag coefficient against the height of the spray and Re number, respectively (based on our experimental data for the spray from the nozzle and formulas (1) and (2)), and also the curves following from formulas (3) and (5). It is seen that the hydrodynamic drag calculated from the experimental data is much (four to seven times) lower than that obtained from formulas (3) and (5), especially at distances $z > 300$ mm from the nozzle, where the experiment gives $C_{d} < 0.2$.

In an attempt to explain such an anomaly in the behavior of the drag coefficient, we put forward a number of hypotheses. These are (1) polydispersity of the spray, which is responsible for the difference in the motion of drops varying in size and for the features appearing in the averaged (integral) motion of the dis-
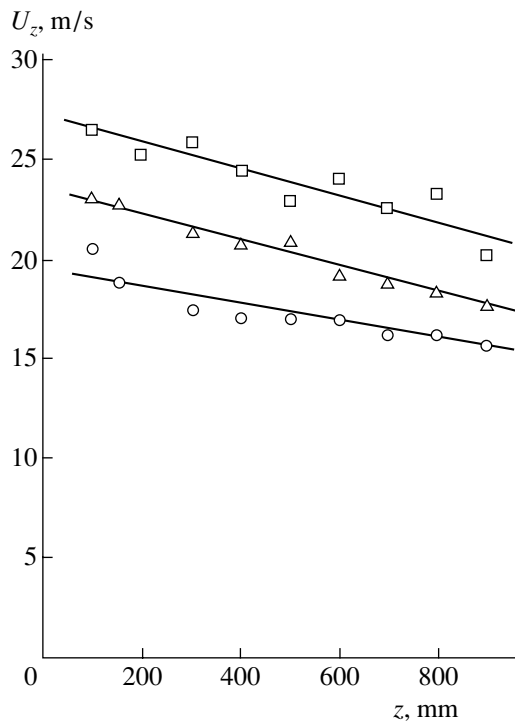
U_z, m/s

Fig. 3. Variation of the drop velocity with the height of the spray. $p = (\bigcirc)$ 300, $(\triangle)$ 500, and $(\square)$ 900 kPa.
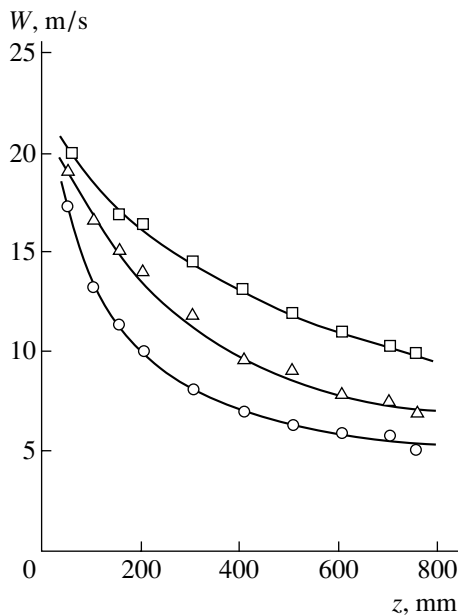
W, m/s

Fig. 4. Variation of the gas velocity with the height of the spray. The symbols mean the same as in Fig. 3.

perse phase; (2) deformation of the drops in the turbulent flow, in particular, oscillation of their shape and, accordingly, variation of their transverse (relative to the flow) size; (3) macroscopic local inhomogeneity of the drop flow structure, which shows up as the group motion of the drops; (4) direct effect of velocity pulsa-

tion in the gas flow about a drop; (5) extension of the Stokes regime for the flow around a drop into the range of transition Reynolds numbers (Re = 1–100) because of the turbulence of the flow incoming to the drops; and (6) the crisis of hydrodynamic drag (a phenomenon that is well known for flow around a rigid sphere at Re ~ 10^5), which is observed even at Re $\geq$ 50 because of the high turbulence of the gas flow.

After these six hypotheses had been quantitatively evaluated based on our experimental data and data available from the literature, the first two collapsed. Indeed, if a polydisperse set of drops is replaced by a monodisperse one with a mean size d_{32} and the same total mass and surface area of the drops, the mean drag coefficient may change by several tens of percent rather than by several times. Deformation, flattening of the drops in the direction of the relative velocity of the phases, may raise but not decrease the drag force and coefficient [8].

Preference was initially given to the third hypothesis, according to which particles of a dispersed liquid in a spray produced by a mechanical nozzle move not individually but largely in the form of clusters, agglomerates, bunches, and clouds [4]. These groups arise when the separate filaments and films of the liquid that form upon disintegration of the jet flowing out of the nozzle decay [15]. It is noteworthy that the volumetric concentration of the liquid inside and between the clusters is, respectively, higher and lower than the mean concentration over the volume, where the amount of the clusters is sufficiently large. Therein lies the meaning of the term "macroscopic local inhomogeneity" of the flow structure of the drop, which takes into account the fact that the clusters are noticeably larger than individual drops. In addition, the gasdynamic drag of a cluster is less than the total drag of the drops moving separately. This explains the new phenomenon discovered, which was initially called not the crisis of drag resistance but merely the drag anomaly. This anomaly showed up as lower experimental values of C_d in Figs. 6 and 7 as compared with the known values obtained from formulas (3) and (5). This hypothesis was indirectly verified in [16], where groups of drops are distinctly seen in photos taken of the drop flow in the spray from the nozzle.

However, the third hypothesis also fails if one takes into consideration the fact that the anomaly (or crisis) of the drag arises at a certain distance from the nozzle (Fig. 6) where the concentration of the drops is appreciably lower than that near the base of the spray cone, where any anomaly is absent. One may therefore expect that the breaking of the group motion of drops due to the turbulent pulsation of the gas flow will intensify with distance from the nozzle.

The fourth hypothesis, the direct effect of gas velocity pulsation, also turns out to be untenable: this effect may increase, rather than decrease, C_d and only by a few percent.

The fifth hypothesis relies on an analogy with liquid flow through a tube: in going from the laminar to the turbulent regime, the drag coefficient reaches a local minimum [2]. Using the concept of the laminar viscous surface layer that borders the turbulent boundary layer from the outside (an idea similar to that used in the theory of near-wall turbulence [1, 2]), we even succeeded in constructing a model that accounts for the extension of the Stokes regime of the flow about a drop into the range of transition Reynolds numbers Re = 50–120 (Figs. 7, 8). However, the construction of the model required that a number of poorly substantiated assumptions be made. Yet this model cannot answer the question why the pseudo-Stokes regime is absent near the spray base, where Reynolds numbers are smaller, and why the drag anomaly emerges only at distances $z >$ 100 mm from the nozzle (Figs. 6, 7).

Thus, only the sixth hypothesis remains to be checked. It states that an early (that is, arising even at Re $\approx$ 50 instead of $\sim 10^5$) crisis of hydrodynamic drag takes place in the spray because of the highly turbulent flow about the drops. Here, the following circumstances are worth noting.

The Re dependence of the drag coefficient for a solid sphere C_d(Re) in the range $10^{-2} < $ Re $< 10^6$ is well known. It was obtained by generalizing a large body of experimental data [1, 3, 8], and its plot (Fig. 8), taken from [1], is sometimes called the Rayleigh curve [9].

This curve can be subdivided into several portions. At Re < 1, the flow is laminar and is described by the Stokes formula $C_d = 24/$Re.

At Re ≈ 20, the laminar boundary layer in the stern region (a polar angle $\theta_l \leq 180°$) separates to form two return-flow vortices of the continuous phase [8, 9].

For $20 <$ Re < 100, the line of separation shifts upstream of the flow up to $\theta_l \approx 120°$ and the vortices behind the sphere grow to 1.2 times the sphere diameter. The drag coefficient exceeds the Stokes value and can be approximated (up to Re = 400–500) by Klyachko formula (3) or by the formula [8, 9, 17]

$$C_d = 18.5/\text{Re}^{3/5}. \qquad (6)$$

For $100 <$ Re < 500, the stern vortices start oscillating and the line of separation of the boundary layer shifts to $\theta_l \approx 80°$ [8, 9].

At Re ≈ 500, the vortices are separated by the flow and drift downstream into the stern wake. At a certain point T downstream from the line of separation in the wake behind the sphere, the transition to turbulent flow occurs [2].

At $500 <$ Re $< 10^5$, the flow about the sphere is usually called turbulent. However, it would be more appropriate to call it a mixed flow, since it is laminar upstream from the line of boundary layer separation ($\theta_l \approx 80°$) and turbulent downstream from the point T. As the Reynolds number grows, the point T moves upstream toward the spherical surface. When this point
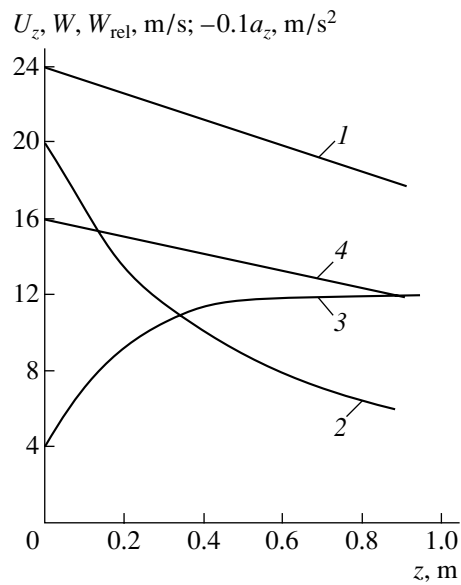

Fig. 5. Approximation of the velocities of the phases and the acceleration of the drops at the axis of the spray: (*1*) velocity U_z of the drops, (*2*) gas velocity W, (*3*) relative velocity of the phases $W_{rel} = U_z - W$, and (*4*) acceleration of the drops taken with opposite sign ($-0.1a_z$) at $p = 500$ kPa.
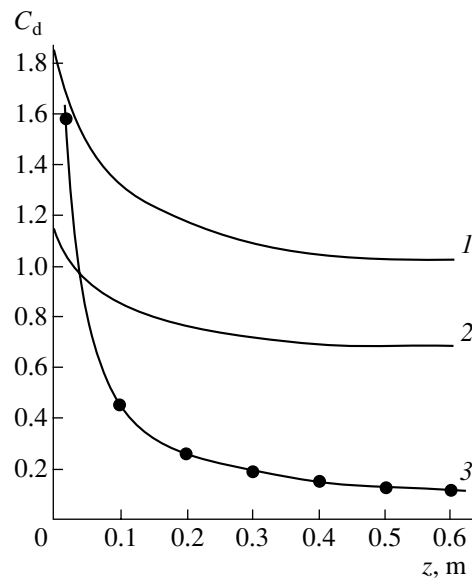

Fig. 6. Variation of the drag coefficient with the height of the spray: (*1*, *2*) calculations by formulas (3) and (5), respectively, and (*3*) experimental data for the spray at $p = 500$ kPa and $d = d_{32} = 140$ μm.

reaches the line of separation ($\theta_l \approx 80°$), the flow throughout the separated layer becomes turbulent and the drag coefficient in this wide region of Re remains almost unchanged ($C_d \approx$ const ≈ 0.5).

At $\text{Re}_{cr} \approx 2.5 \times 10^5$, we come up against the well-known crisis of hydrodynamic drag for a sphere: the drag coefficient C_d decreases drastically down to 0.1–
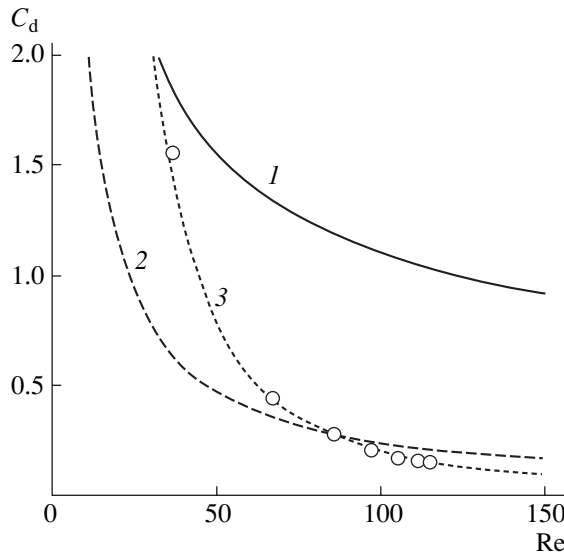
Fig. 7. Dependence $C_d(\text{Re})$ obtained by calculation with (*1*) formula (3) and (*2*) the Stokes formula. (*3*) Approximation $C_d = 2000/\text{Re}^2$ of the experimental data ($\bigcirc$) for the spray.
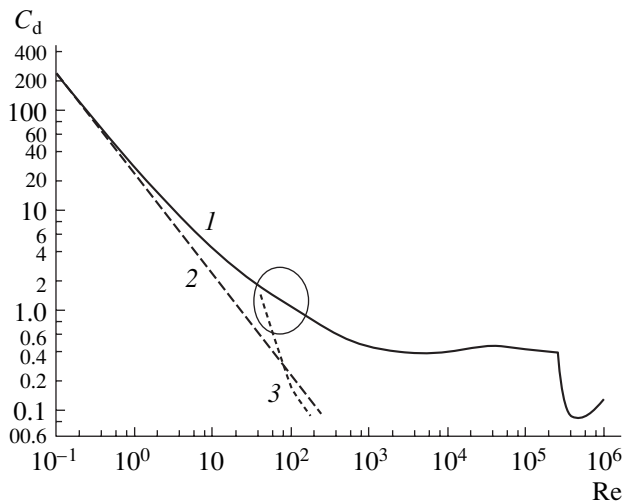

Fig. 8. Drag coefficient of the sphere as a function of the Reynolds number: (*1*) Rayleigh curve [1], (*2*) the Stokes theory (the circle outlines the region in the spray that is critical in Reynolds number), and (*3*) decrease in C_d in the region marked (experiment [4]).

0.2 [1–3, 8]. The onset of the crisis is accompanied by the separation of the laminar boundary layer, which becomes turbulent. Simultaneously, the line of separation shifts downstream toward $\theta = 120°–140°$, that is, downstream from the transition point T at $\theta_l \approx 100°$. As this takes place, the flow about the sphere approaches the idealized flow pattern and the pressure in the stern region of the sphere increases markedly, reducing the total hydrodynamic drag [2].

The description above refers to the case when the flow about a sphere is initially laminar. It is known that

"the turbulence of the incoming flow affects the crisis of drag resistance. The higher the turbulence, the smaller the Re at which the boundary layer turbulization occurs. As a result, the decrease in the drag coefficient starts at smaller Reynolds numbers (and is observed in a wider interval of Re) [3]. It was reported [2] that, as the degree of turbulence $\varepsilon = w'_m/\overline{w}$ (where w'_m is the velocity pulsation amplitude and $\overline{w}$ is the averaged gas velocity) increases from 0.5 to 2.5%, the critical Reynolds number Re_{cr} decreases from 2.70×10^5 to 1.25×10^5 (i.e., by half and not by three orders of magnitude!). In [2], it is also noted that "the crisis of drag resistance may occur at Reynolds numbers that are considerably smaller than the critical value if the boundary layer is turbulized artificially...."

It turned out that, when the turbulence of the flow is high ($\varepsilon \approx 30\%$ or higher, as in a spray produced by a nozzle [4]), the crisis of drag resistance may occur at transition Reynolds numbers as low as $\text{Re} \geq 50$, as follows from the experiment.

The early crisis of hydrodynamic drag for drops in the spray explains a number of features of the two-phase flow that were detected in the experiment. These are the following: (1) even at a distance of 1 m from the nozzle, the relative velocity of the phases is considerable, reaching 12 m/s (Figs. 1, 2, 5); (2) the momentum transferred from the liquid to the gas both in the free spray and in the sprayer is roughly half the initial momentum of the liquid jet; (3) the crisis arises at a certain distance from the nozzle rather than in the immediate vicinity of it. Therefore, two flow regions can be distinguished in the spray: (i) the base region, where the phases interact extensively, exchanging momentum, and (ii) the self-similar region at distances $z \geq 300$ mm, where interaction between the phases is much weaker, their momentum fluxes are almost invariable, and the radial profiles of all the hydrodynamic characteristics are self-similar [4].

Numerical simulations of the two-phase flow in the spray in the two-dimensional model based on formulas (3) or (5) (i.e., without allowance for the early crisis) invariably failed: the gas velocity turned out to be much higher and the velocity of the liquid, much lower than in experiments [4]. Conversely, when experimental values of C_d including the early crisis were used in calculating the force of interfacial interaction, agreement between calculated and experimental velocities of the phases was greatly improved [18].

Thus, we conclude that highly turbulized flows like those in sprays produced by nozzles may exhibit an early crisis of hydrodynamic drag for disperse phase particles even at transition Reynolds numbers as low as several tens and not only at $\sim 10^5$, as was considered earlier.

REFERENCES

1. H. Schlichting, *Boundary Layer Theory* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1974).

2. G. G. Loĭtsyanskiĭ, *Mechanics of Liquids and Gases* (Nauka, Moscow, 1978).

3. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).

4. N. N. Simakov, Candidate's Dissertation (Yaroslavsk. Pedagog. Inst., Yaroslavl, 1987).

5. K. S. Shifrin and V. I. Golikov, in *Proceedings of the VI Interdisciplinary Conference on the Physics of Clouds, Precipitations, and Atmospheric Electricity* (Akad. Nauk SSSR, Moscow, 1961), pp. 266–277.

6. S. V. Zhigulev, Uch. Zap. Tsentr. Aérogidrodin. Inst. **13** (5), 142 (1982).

7. O. M. Belotserkovskiĭ and Yu. M. Davydov, *Coarse Particle Method in Gas Dynamics* (Nauka, Moscow, 1982).

8. R. I. Nigmatulin, *Dynamics of Multiphase Media* (Nauka, Moscow, 1987), Part 1.

9. B. I. Brounshteĭn and G. A. Fishbeĭn, *Fluid Dynamics and Mass-and-Heat Exchange in Dispersive Media* (Khimiya, Leningrad, 1977).

10. S. Ergun, Chem. Eng. Prog. **8** (2), 89 (1952).

11. N. I. Gel'perin, B. N. Basargin, and Yu. G. Zvezdin, Teor. Osn. Khim. Tekhnol. **6**, 434 (1972).

12. N. I. Gel'perin, B. N. Basargin, Yu. G. Zvezdin, and V. V. Vlasov, Teor. Osn. Khim. Tekhnol. **8**, 463 (1974).

13. Yu. G. Zvezdin and B. N. Basargin, Teor. Osn. Khim. Tekhnol. **16**, 715 (1982).

14. Yu. G. Zvezdin, N. N. Simakov, A. P. Plastinin, and B. N. Basargin, Teor. Osn. Khim. Tekhnol. **19**, 354 (1985).

15. V. A. Borodin *et al.*, *Spraying of Liquids* (Mashinostroenie, Moscow, 1967).

16. B. N. Basargin, V. I. Katalov, and Yu. I. Gushchin, *Mass-and-Heat Exchange Processes in Chemical Technology: Collection of Scientific Articles* (Yaroslavsk. Politekhn. Inst., Yaroslavl, 1976), pp. 174–177.

17. R. B. Bird, W. E. Stewart, and E. N. Lightfoot, *Transport Phenomena* (Wiley, New York, 1960; Mir, Moscow, 1974).

18. N. N. Simakov, Izv. Vyssh. Uchebn. Zaved., Khimiya Khim. Tekhnol. **45** (7), 125 (2002).

Translated by N. Mende

GASES
AND LIQUIDS

Mechanism of Turbulent Pulsation in Channels

L. N. Pyatnitsky

Joint Institute of High Temperatures, Russian Academy of Sciences, Moscow, 127412 Russia
e-mail: pyat7@mail.ru
Received July 18, 2003

Abstract—The space–time pulsation field of hydrodynamic parameters for channel turbulent flow is derived from the wave model of turbulence. Conditions for pulsation field randomization are analyzed. Calculations of the longitudinal and transverse pulsations are compared with the Reichardt's measurements. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The form of an action that disturbs a turbulent flow and the mechanism behind turbulent pulsation were discussed in [1, 2]. The formation of a turbulent boundary layer and laminar sublayer was described in [3]. The pulsation of hydrodynamic parameters results from a superposition of disturbances that arise in the field of a high flow-velocity gradient at the channel wall and propagate in the form of wave packets.

The velocity gradient is a result of flow-on-wall rubbing. The thickness of the packet is comparable to the amount of the primary disturbance, and the hydrodynamic parameters within the packet vary (pulsate) with a spatial period equal to its thickness. In the flow core, near-spherical wave packets superpose. At the walls, where the velocity gradient is high, the wave front of the packets is broken. The region of front breaking is considered as the boundary layer.

In this work, we (i) derive relationships for the space–time pulsation of hydrodynamic parameters in the field of a velocity gradient; (ii) suggest a method for determining the thicknesses of the boundary layer, laminar sublayer, and vortex formation region; (iii) consider the transformation of regular disturbances of the parameters into irregular ones; and (iv) analyze the properties of the random pulsation field. Calculations for the velocity field pulsation in a parallel-plate channel are compared with the Reichardt's measurements. An interpretation of the experimental dependences of the longitudinal and transverse components of pulsation on the distance to the wall is given.

SPACE–TIME PULSATION FIELD

A primary disturbance can be represented [3] by a region of overpressure $\Delta p = p_0 f(r)$ that is located at the wall and is bounded by a semisphere of radius a ($r \leq a$). This disturbance propagates with the velocity of sound c in the form of a wave packet of thickness $2a$. Within the packet, hydrodynamic parameters pulsate. The pul-

sation of the velocity u, pressure p, density ρ, and other parameters is defined through a function $f(r)$, and the pulsation amplitude is inversely proportional to the distance traveled by the wave (see, e.g. [4]):

$$u, p, \rho \sim f\!\left(\frac{l_0 - ct}{a}\right)\!\frac{1}{ct} \quad \text{for} \quad |l_0 - ct| \leq a,$$
$$u, p, \rho = 0 \quad \text{for} \quad |l_0 - ct| > a. \tag{1}$$

Here, l_0 is the distance to the point of observation and t is the time. In an unbounded quiescent medium, the wave retains its shape. In a channel, however, the wave front is distorted because of reflections from the walls and a complex velocity profile in the channel cross section. Let us track a sample point of the wave that propagates in the field of the flow velocity $\mathbf{U}$. The propagation direction of the wave at this point (the propagation direction of an acoustic beam in geometrical acoustics) is given by a unit vector $\mathbf{s}$.

In the field of the velocity $\mathbf{U}$, the direction of the vector $\mathbf{s}$ at the sample point is found from the equation [4]

$$\frac{d\mathbf{s}}{dl} = \frac{1}{c}\text{curl}\,\mathbf{U} \times \mathbf{s}, \tag{2}$$

where dl is an element of the path covered by the point.

If $\mathbf{U} = \text{const}$, the sample point moves along a straight line and the wave remains spherically symmetric. If $\mathbf{U}$ varies over the channel cross section, the vector $\mathbf{s}$ deviates from the initial direction and the wave front deforms.

Let the flow be directed along the x axis of a parallel-plate channel of height $z = d$. Our aim is to estimate the total angle of rotation of the vector $\mathbf{s}$ in the velocity field $U(z)$. Introducing the polar, ϑ, and azimuthal, φ, direction angles for the vector $\mathbf{s}$, changing the velocity $U(z)$ to the Mach number $M(z)$, and integrating (2), we

arrive at ($M \ll 1$)

$$\sin\vartheta\cos\varphi = \sin\vartheta_0\cos\varphi_0 + M(z);$$
$$\sin\vartheta\sin\varphi = \sin\vartheta_0\sin\varphi_0, \tag{3}$$

where ϑ_0 and φ_0 are the initial values of the angles.

A solution to set (3) gives the direction angles $\{\vartheta, \varphi\}$ of the vector **s** as functions of its initial orientation $\{\vartheta_0, \varphi_0\}$, coordinate z, and given velocity distribution $M(z)$:

$$\sin\vartheta = \sqrt{\sin^2\vartheta_0 + 2M(z)\sin\vartheta_0\cos\varphi_0 + M^2(z)},$$
$$\sin\varphi = \frac{\sin\vartheta_0\sin\varphi_0}{\sqrt{\sin^2\vartheta_0 + 2M(z)\sin\vartheta_0\cos\varphi_0 + M^2(z)}}. \tag{4}$$

Knowing the orientation $\{\vartheta, \varphi\}$ of the vector **s**, one can calculate the trajectory of the sample point in the channel. First, we will find the coordinates of this point $\{\xi, \eta, \zeta\}$ vs. the distance l traveled by the point in the half-space, i.e., in the absence of the upper wall. According to (4), the position of a point relative to the center of the wave is defined by the set of parametric differential equations

$$\frac{d\xi}{dl} = \sin\vartheta_0\cos\varphi_0 + M(z),$$
$$\frac{d\zeta}{dl} = \sqrt{\cos^2\vartheta_0 - 2M(z)\sin\vartheta_0\cos\varphi_0 - M^2(z)}, \tag{5}$$

$$\xi^2 + \eta^2 + \zeta^2 > a^2.$$

For the wave in the channel, it is taken into account that a disturbance may occur at an arbitrary point **rp**$\{xp, yp, zp\}$ of either wall with a delay tp relative to $t = 0$. We move on to dimensional parameters using the height d as a spatial scale and the ratio d/c as a time scale. The number n of reflections of the wave from the walls may be expressed through the ordinate ζ : $n = \mathrm{In}(\zeta)$ [3]. Then, set (5) for the coordinates $\{x, y, z\}$ of the sample point in the channel with allowance for reflections takes the form

$$x = xp + \int_0^{l-tp} (\sin\vartheta_0\cos\varphi_0 + M(z))dl,$$

$$y = yp + \sin\vartheta_0\sin\varphi_0(l - tp),$$

$$z = zp + \int_0^{l-tp} \left(\sqrt{\cos^2\vartheta_0 - 2M(z)\sin\vartheta_0\cos\varphi_0 - M^2(z)} \right.$$
$$\left. - 2\mathrm{In}\left[\frac{n+1}{2}(-1)^{n+zp}\right]\right)dl, \tag{6}$$

$$n = \mathrm{In}[(l - tp)\sqrt{\cos^2\vartheta_0 - 2M(z)\sin\vartheta_0\cos\varphi_0 - M^2(z)}],$$

$$(x - xp)^2 + (x - xp)^2 + (x - xp)^2 > a^2.$$

Equations (6) determine the position of the sample point as a function of the distance l traveled by the wave and include the point, **rp**$\{xp, yp, zp\}$, and time, tp, of occurrence of a disturbance with a radius a. The parameter pulsation vector at the sample point is found by jointly solving Eqs. (6) and (1). In this case, it should be taken into account that the distance l_0 to the point of observation changes to l_n because of reflections and the amplitude of the wave decreases $(1 - \alpha)^n$ times; that is, at $|l_n - ct| \le a$, we have

$$u, p, \rho \sim (1 - \alpha)^n f\left(\frac{l_n - ct}{a}\right)\frac{1}{ct}. \tag{1a}$$

The wave front configuration and the pulsation vector field are found by varying the direction angles within the intervals $\varphi_0 \in [0, 2\pi]$ and $\vartheta_0 \in [0, \pi/2]$. The pulsation field evolution is found by varying l so that the variables $\{x, y, z\}$ meet the channel profile.

If the parameters **rp** and tp are sequences, Eqs. (6) describe the evolution of the disturbance wave front structure in a channel. Jointly solving Eqs. (6) and (1a) yields a space–time pulsation field for the hydrodynamic parameters of a channel flow. Software tools currently available, such as Mathematica 4, allow the calculation of this field.

WAVES IN THE FIELD OF A HIGH FLOW-VELOCITY GRADIENT

Let the flow velocity profile in the channel be represented by the formula

$$M(z) = M_0[4z(1 - z)]^{1/m}, \tag{7}$$

where m is a parameter that characterizes the flow regime: $m = 1$ for a laminar flow, $m \gg 1$ (e.g., $m = 5$) for a turbulent flow, and $m = \infty$ for a flow with a cross-section-invariable velocity.

As follows from (4), the trajectory of the vector **s** in the field $M(z)$ departs from a straight line. The deflection depends on the distance l the vector **s** travels in a δ-thick layer with a high velocity gradient.

Differentiation of formula (7) yields an expression for the flow velocity gradient:

$$\frac{d}{dz}M(z) = M_0\frac{\sqrt[m]{4}}{m}\frac{1 - 2z}{(z(1 - z))^{\frac{m-1}{m}}}. \tag{8}$$

In the laminar regime ($m = 1$), this gradient is weak and the deflection of the trajectory is insignificant. In the turbulent regime, the velocity gradient is almost absent in the flow core but grows drastically in a thin near-wall layer of thickness δ. For example, in the turbulent flow with $m = 5$ and $M_0 \approx 0.15$, the gradient at the boundary $z = 0.03$ of this layer increases 30-fold compared with the laminar case ($m = 1$, $M_0 \approx 0.005$).

To visualize the results, consider the evolution of acoustic beams in the section $y = 0$, where $\varphi = 0$ and the bend of the beams is the most noticeable. Under these conditions, relationships (4) are simplified and the orientation of the vector **s** is found from (4) and (7) as

$$\vartheta = \arcsin(\sin\vartheta_0 + M_0[4z(1-z)]^{1/m}). \qquad (9)$$

For large ϑ_0 near the wall, the trajectory length in the layer δ increases rapidly with angle: $l_0 = \delta/\cos\vartheta_0$. Since the trajectory bends, its actual length is much greater: $l \gg l_0$. For angles close to $\pi/2$, there exists an angle ϑ_0 at which the trajectory moves away from the wall by a distance not exceeding a certain value z_*. Putting $\vartheta = \pi/2$, we find the dependence of z_* on ϑ_0 and flow parameters (the minus sign before the radical corresponds to the lower wall):

$$z_* = \frac{1}{2}\left(1 - \sqrt{1 - \left(\frac{1 - \sin\vartheta_*}{M_0}\right)^m}\right). \qquad (10)$$

As was shown [3], in the laminar regime, the trajectories of the vector **s** appear as straight lines almost at any ϑ_0. Moreover, with $m = 1$ and $m = \infty$, they nearly coincide. In the turbulent regime, the situation changes. With $\vartheta_0 < \vartheta_*$, z_* is imaginary, the trajectories are straight lines, and the wave front remains spherically symmetric. However, at $\vartheta_0 \geq \vartheta_*$, z_* becomes real and the shape of the trajectories changes. We will discuss this case in greater detail, since it is of practical importance.

If z_* is real, the sample point moves along the ascending branch, reaches a maximum at $z = z_*$ ($\vartheta = \pi/2$), and returns to the wall, thus describing an arc. Since this arc is symmetric about the line $z = z_*$ and the angle of incidence $|-\vartheta_0|$ equals the angle of reflection ϑ_0 for this specific ray, the motion takes a cyclic character and persists until the wave decays completely.
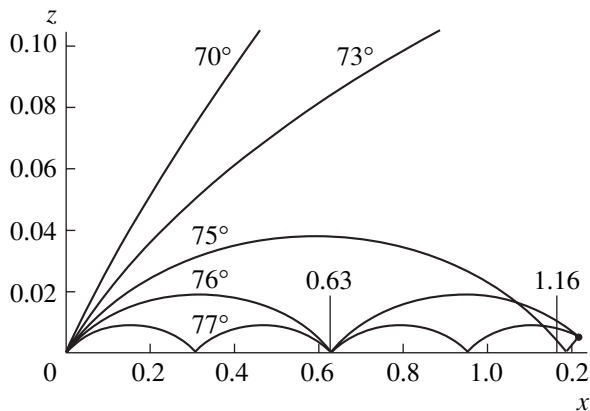

Fig. 1. Propagation of acoustic rays at the wall for $\vartheta_0 = 70°$, $73°$, $75°$, $76°$, and $77°$.

Consider the trajectories with ϑ_0 that are large yet distant from $\pi/2$. Let the wave propagate in the turbulent flow with $M_0 = 0.05$ and $m = 5$. For simplicity and better clarity, we will describe the essence of the process for the near-wall layer height within the interval $z \in [0, 0.1]$ and for the wave travel $l \approx 1.2$. The initial directions of rays are defined by $\vartheta_0 = 70°, 73°, 75°, 76°$, and $77°$ (see Fig. 1).

As follows from Fig. 1, the trajectory with $\vartheta_0 = 70°$ penetrates into the flow core and its direction remains practically unchanged. Theoretically, the ray with $\vartheta_0 = 73°$ describes a cyclic trajectory with $l \approx 4.5$ and $z_* \approx 0.15$. However, it is obvious that, at such l, the disturbance decays without reaching the end of even one cycle. Thus, three trajectories with $\vartheta_0 = 75°, 76°$, and $77°$ are left. Their behavior merits detailed consideration.

First, it should be noted that the trajectories execute a cyclic motion with two to four cycles depending on the angle ϑ_0 over the length $l = 1.2$. Therefore, they intersect with each other. For example, in the section $x = 1.16$, the rays appear in order of angles $75°–77°–76°$ instead of the expected $77°–76°–75°$. In the wave, the angle ϑ_0 certainly varies continuously, so that the rays mix up completely in a layer of height $z_*(75°) \approx 0.04$ at $\vartheta_0 \geq 75°$.

In the mixing area, as in the flow core, the wave interaction forms its pulsation field. However, here the field changes because of the decay of the wave packet front (acoustic ray mixing).

Since the mechanism of wave interaction in this area is different from the rest of the space–time field, this area should be considered separately and may be treated as a boundary layer. Its inner interface (with the flow core) contains trajectories along which the disturbance does not decay at least within a cycle. In Fig. 1, this interface is determined by $z_*(\vartheta_0 = 75°) \approx 0.04$.

According to Fig. 1 ($m = 5$, $M_0 = 0.05$), trajectories with $\vartheta_0 < 75°$ leave the boundary layer for the core through the inner interface. However, in a narrow range of ϑ_0 in going from the boundary layer to the core, the path in the cycle is so long that the disturbance decays without reaching the wall, although formally (theoretically) the trajectories close the cycle (reach the wall).

Let a 100-fold attenuation of the wave amplitude be the test for decay. Then, a disturbance with an initial size $2a = 0.1$ decays over the length $l = 5$. Figure 2 shows solutions to set (5) for $l = 5$ and $\vartheta_0 \in [72°, 73°]$ with step $\Delta\vartheta_0 = 0.1°$ that depict vortical portions of the trajectories.

However, Fig. 2 gives a simplified picture of vortex formation, showing only the set of trajectories in the plane $\varphi = 0$. Outside this plane, the longitudinal component of the wave velocity depends on $\cos\varphi$ and drops, becoming a variable quantity. Simultaneously, the curvature of the trajectories grows and high-curva-

ture paths appear. Note that despite the condition $M_0 \ll 1$, the Mach number is the basic parameter governing the shape of vortical trajectories.

At the outer interface (within a narrow layer in the immediate vicinity of the wall), another feature of the turbulent flow shows up. It is known that the velocity pulsation can be related to the specific energy flux $c\rho u^2$. When the wave packet reflects from the wall, the velocity pulsation at the wall vanishes and the pulsation energy is transferred to pressure and density pulsations. Because of the absence of the velocity pulsation, this layer next to the wall can be considered as a laminar sublayer of the turbulent boundary layer. The thickness δ_L of this sublayer depends on the hydrodynamic parameter distribution in the primary disturbance, which is represented by the function $f(r, t \leq a)$, and, as is easy to check, varies within $\delta_L = (0.1–0.2)a$, depending on the form of this function.

One more consequence of ray interweaving in the boundary layer is the coincidence of the point (zones) of reflection for rays with different ϑ_0. One such point in Fig. 1 is near the coordinate $x = 0.63$. When the packet reflects, the incident and reflected waves interact. In the interaction domain ($z \leq a$), the pulsation amplitude doubles and simultaneously the pulsation frequency is halved. The same is true for density pulsation. However, the pulsations of the parameters in adjacent sectors of one wave are coherent. Therefore, the coincidence of the zones of reflection causes nonlinear effects and spatial angular modulation rather than mere superposition.

Under such conditions, zones of height a where elementary waves coming at different angles ϑ_0 reflect simultaneously serve as new sources of primary disturbances. Since the waves interact in the region where rays mix up (this region is the basic sign of the boundary layer), a primary disturbance may be assigned the size $a = z_*(\vartheta_*)$.

Thus, analyzing physical processes under the conditions of high flow-velocity gradients, we revealed the mechanisms of formation of singularity regions and defined them in physical terms (boundary layer, laminar sublayer, and vorticity area in the turbulent flow). The numerical values of these parameters depend on the travel l of the wave. It, in turn, is related to the minimal number of pulsations that is necessary for a specific problem to be solved.

PULSATION FIELD RANDOMIZATION

A solution to the linear (!) wave equation yields the wave front position and pulsation in space and time (the space–time wave field). Evidently, this field is determinate and regular if one wave propagates in space. In a channel, this field includes disturbances generated by reflections from the walls.
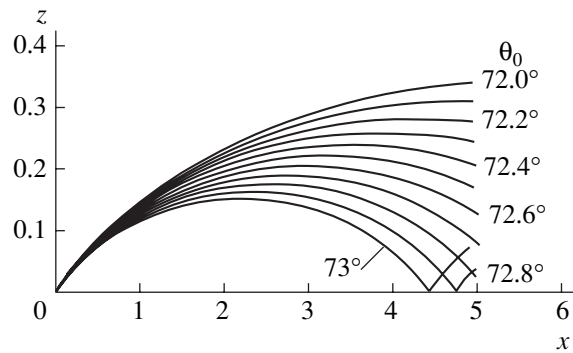
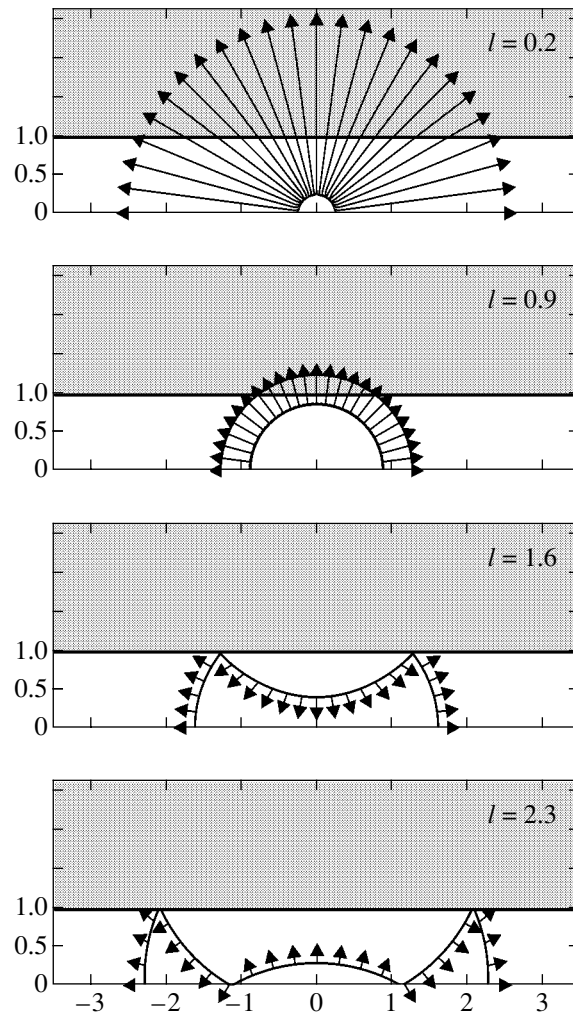**Fig. 2.** Trajectories adjacent to the boundary layer.

Fig. 3. Wave front and velocity pulsation in the plane $y = 0$ for various l ($a = 0.03$).

Figure 3 shows the variation of the velocity pulsation vector field with wave front position for the wave propagating over the plane $y = 0$ of the flow core in a plane-parallel channel. The channel is shown as a bright (white) fringe; the outside of the channel is
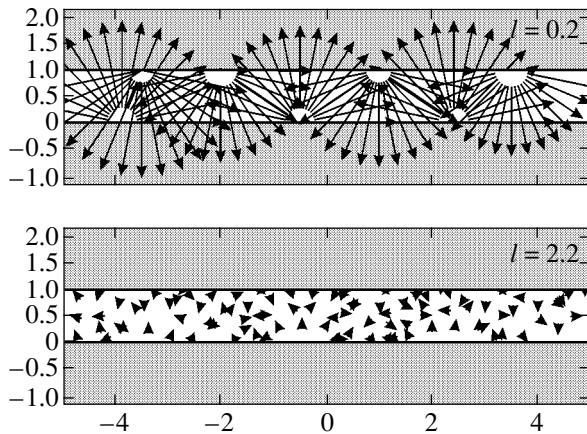
Fig. 4. Seven waves and their associated velocity pulsations in the channel for $l = 0.2$ and 2.2.

shaded. The travel l of the wave is shown at the upper right of the panels. The velocity vector magnitude (length) corresponds to the pulsation amplitude. As the wave propagates, the velocity pulsation vector changes its direction according to (5) and the pulsation amplitude decreases smoothly in accordance with (1a). However, the space–time field remains determinate and regular, as in free space.

When many waves propagate in the channel, to trace the history of each wave is a challenge, although their origination and evolution follow particular rules and, as follows from the above, can be calculated. The randomization of the space–time velocity pulsation field is demonstrated in Fig. 4 with seven disturbances propagating in the section $y = 0$ of the channel. The pulsation field is shown for $l = 0.2$ and 2.2. As before, the arrows indicate the directions of the velocity pulsation vector for $a = 0.03$, and the magnitudes (lengths) of the vectors are the pulsation amplitudes at given points of the wave front. The wave fronts are not shown.

At the initial state of propagation of the seven waves ($l = 0.2$), the disturbance field remains fairly regular. However, at $l = 2.2$, the pulsations appear chaotic. Note, however, that the vector field is constructed for a spread in disturbance origination times $tp \leq 0.2$; i.e., the waves emerge almost simultaneously (this is needed for better visualization of the vectors). Therefore, randomization is observed only for pulsation directions. Actually, the range of tp is unlimited and, when disturbances arise in sequence, the pulsations differ both in direction and amplitude (cf. the upper and lower panels in Fig. 3).

To gain a better insight into the process, let us analyze pulsations in the flow core in a plane-parallel channel. Expression (1a) for the velocity in a wave packet is now replaced by

$$u = (1-\alpha)^n \cos\left(\frac{\pi}{2}\frac{l_n - t}{a}\right)\frac{l_n - t}{t}, \qquad (1b)$$

where $a = 0.1$ and $\alpha = 0.1$.

The numerical factor in (1b) was set equal to unity, since we are interested primarily in velocity pulsations.

With many waves propagating in a channel, one should sum disturbances arising at different places and time instants. Here, the stage of pulsation evolution is of great concern. We will concentrate on steady-state conditions. If k disturbances with a lifetime τ are generated in the flow within a time of observation T, $kn\tau/T$ reflected waves will be permanently present in the channel on average. The inverse value of this parameter, β, characterizes the extent to which the process is resolved in time or space.

The problem of finding the setup time for the pulsation structure is akin to the problem of finding the liquid elevation in a vessel in the case of constant delivery and elevation-dependent discharge. Wave generation–decay balance will be described by an expression proportional to $\exp(-t/\tau)$. The regime is considered steady starting with $t_1 = 3\tau$, when the number of waves in the channel differs from the equilibrium value by no greater than 5%.

To characterize the space–time pulsation field, either pulsation is recorded at a fixed point in the channel or the disturbance distribution in a given direction is visualized with streak photography. Consider first velocity pulsations at a given point. When they are in dynamic equilibrium, the time of observation must meet the condition $t \in [t_1, t_2]$, where $t_1 = 3\tau$ and $t_2 \leq T$. Also, the channel volume under study must be large enough so as to provide dynamic equilibrium between pulsations at the point of observation. The latter condition can be satisfied by taking into account only the disturbances offset from the point of observation by a distance that is equivalent to the wave lifetime τ. Then, for $\tau = 10$, we find the boundaries of the volume $\{[x_1, x_2], [y_1, y_2], (z_1, z_2)\} \in \{[0, 20], [0, 20], (0, 1)\}$ centered at the point $x = y = 10$. The pulsation at a point $\{x, y, z\}$ is found from (6) and (1b) by summing the piecewise smooth function $u(x_1, x_2, y_1, y_2, z_1, z_2, t_1, t_2, xp, yp, zp, tp, x, y, z, n, k, \tau, T)$ over all k disturbances and their reflections from the walls.

Consider the case $k = 100$. The point of observation is specified as $\{x, y, z\} = \{10, 10, 0.7\}$, the interval of observation is $[t_1, t_2] = [30, 50]$, and the process duration $T = 50$. With $a = \alpha = 0.1$ and $\tau = 10$, the pulsation amplitude diminishes roughly 300-fold within this lifetime. Pulsations of lesser amplitudes may be neglected in the context of our consideration. When the pulsations were numerically calculated (during observation of the process), the process duration T was divided into $N = 2000$ subranges. It was also necessary to choose the form of the sequences **rp** and tp. The only versions that are meaningful are those where disturbances occur at random or fixed points **rp**, while time instants tp may vary randomly, periodically, or quasi-periodically. The velocity pulsation and pulsation spectra for these three sequences **rp** and tp are shown in Fig. 5 (the upper and lower panels depict, respectively, the time dependence
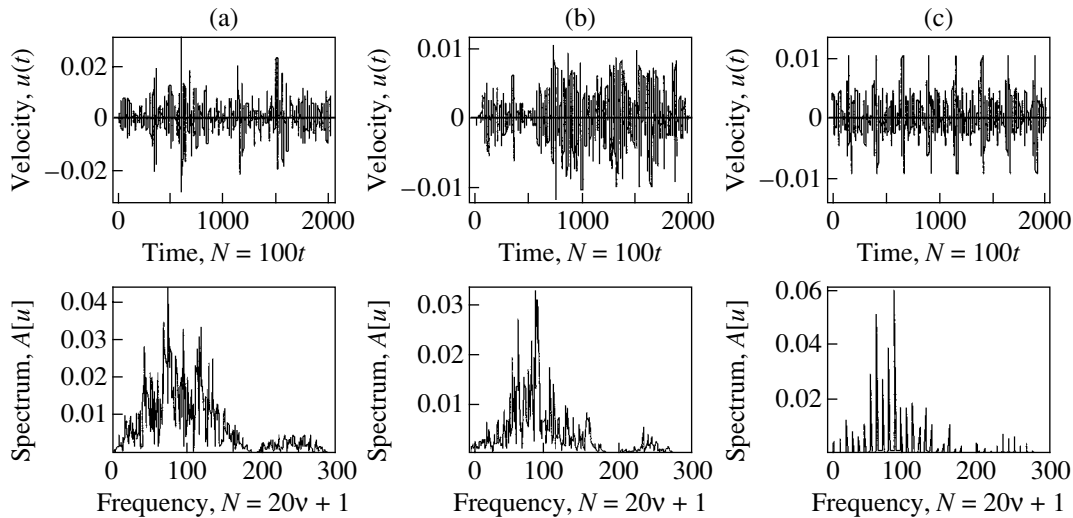
Fig. 5. Velocity pulsations and spectra at the point $x = y = 10$, $z = 0.7$ within the interval $t \in [30, 50]$: (a) random **rp** and tp, (b) five given **rp** and random tp, and (c) five points **rp** with a period **rp** $= T/20$.

of the velocity $u(t)$ and velocity pulsation spectra in the form of a function of frequency $A(\nu)$). The time and frequency are represented through the number of intervals N as $N = 100t$ and $N = 20\nu + 1$. Fourier transformation halves the number of resolvable spectral elements. Figure 5a is constructed for random sequences **rp** and tp, Fig. 5b corresponds to the version where disturbances arise at five regularly arranged points at random times, and Fig. 5c differs from Fig. 5b in that the sequence tp is also regular and has a period of $0.05T$.

Let us find the spatial distribution of the velocity and the spatial spectrum along the x axis in the volume {0–20, 0–20, 0–1} at the time $t = T = 50$, all other things being the same. The dependence of the pulsation on the coordinates r of the segment is obtained by summing the piecewise smooth function $u(x_1, x_2, y_1, y_2, z_1, z_2, t_1, t_2, xp, yp, zp, tp, x, y, z, r, n, k, \tau, T)$ over k waves for the sequence within this segment. However, to keep a constant disturbance concentration over the segment, it is necessary to take into account the edge effect. To do this, we extend the region of disturbance origination on both sides by a distance corresponding to the wave lifetime and correct the number of disturbances. For $\tau = r = 10$, this gives $k = 200$ and the working volume $\{x_1 - x_2, y_1 - y_2, z_1 - z_2\} \in \{0 - 40, 0 - 20, 0 - 1\}$. The pulsation structure is shown in Fig. 6.

Here, the velocity pulsation is plotted against the length of the segment for the interval $x \in [10, 30]$ with a step $\delta x = 0.01$ at $y = 10$, $z = 0.7$, and $T = 50$. The velocity distributions $u(x)$ are shown in the upper panels; the corresponding spectra, in the lower panels. The length and spatial frequency are represented by the number N of resolvable elements as $N = 100x$ and $N = 20\nu + 1$. The resolution β corresponds to the spectral range $N = 1000$ (the frequency bandwidth $\nu = 50$); however, only the most informative part ($\nu \approx 15$) of the spectra are

shown. The properties of the sequences **rp** and tp in Figs. 6a–6c and 5a–5c, respectively, are identical.

From Figs. 5 and 6, it follows that the behavior of the pulsations in space and time are similar. The maxima of the envelopes are at $\nu = 1/2a$. Individual spectral lines can be distinguished. The first line is at a frequency $\nu = 2$, which coincides with the mean rate of disturbance origination. However, there are differences. One is that, in the low-frequency (long-wave) range, the density of the spatial spectra far exceeds the density of the time spectra. This difference is easy to explain.

At a given point, the characteristic frequency is related to the thickness $2a$ of the wave packet. At other frequencies, the density of the spectrum depends on the coherence length of a train of pulsations superposed on one another. At the point of observation, the possibility that the train length will exceed $2a$ decreases and, accordingly, the spectrum becomes sparser at low frequencies (Fig. 5). When the wave crosses the segment where the spatial spectrum is studied, the coherence length increases with wave packet radius automatically.

From Figs. 5 and 6, another feature of the pulsations is seen: in both figures, panels (a)–(c) differ in spectral line contrast, i.e., in pulsation regularity. In general, the irregularity of a wave field may be caused by wave equation nonlinearity, the shape of the channel boundary, or the boundary values of the field. That is, the wave field irregularity is a function of disturbance sources. Since we solve the linear wave equation, the irregularity may be associated with the disturbance sequences **rp** and tp or with the channel boundaries that have the form of Sinay billiards, where the wave front curvature plays the role of curved boundaries.

An irregular wave field may be stochastic or random. The realizations of stochastic and random processes (in our case field waveforms) are identical. Both
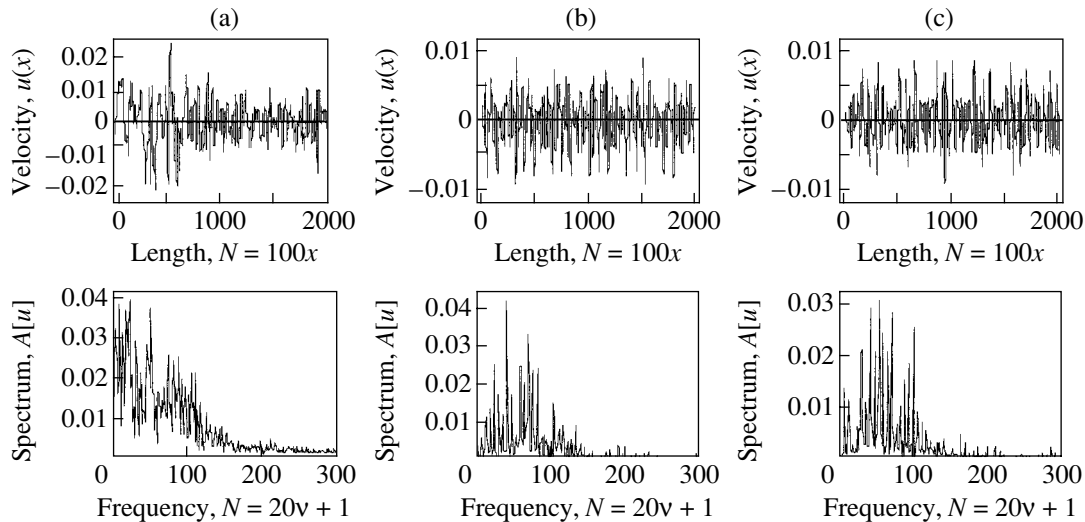
Fig. 6. Pulsation structure along the line $x \in [10, 30]$, $y = 10$, $z = 0.7$ at the time instant $T = 50$: (a) random **rp** and tp, (b) five given **rp** and random tp, and (c) five points **rp** with a period $tp = T/20$.

have a continuous Fourier spectrum and descending autocorrelation function. However, the former is generated by a system with a finite number of degrees of freedom.

The sequences **rp** and tp affect the pulsation properties, as noted above. If disturbances in both sequences are distributed randomly (Figs. 5a, 6a), the pulsation exhibits a continuous spectrum with sharp peaks. For regular **rp** and random tp sequences (Figs. 5b, 6b), the half-height width of spectral lines shrinks and the spacing between the peaks increases noticeably. When both sequences are regular, individual peaks dominate in the spectrum (the spectrum in Fig. 5c is virtually a line spectrum).

In Fig. 5c, the spectrum consists of 23 lines with an average frequency separation of 0.4. The separation depends on the rate of origination of disturbance groups (each consisting of five disturbances), which make up a periodic sequence **rp**. This spectrum describes the velocity periodic variation, which may be considered regular at the point of observation. It is of interest that the continuous part of the associated spatial spectrum (Fig. 6c) is more pronounced and the pulsation along the x axis is more likely to be stochastic than regular.

Thus, the pulsation regularity depends on the sequences **rp** and tp. Two ultimate states of pulsation exist. With primary disturbances originating randomly, the pulsation at the point of observation is also random. When disturbances are generated in a regular manner, the velocity pulsation at the point of observation in a given volume becomes regular. However, the fast variation of the pulsation field, which is due to the high acoustic wave velocity and the variety of disturbance configurations, gives the impression that the pulsation is random even if it is actually regular. Note that the space–time pulsation field can be constructed from a given set of process parameters. On the other hand, the strong dependence of the spectra on initial conditions allows for the solution of the inverse problem: to find these conditions from the spectra available.

PULSATION COMPONENT REDISTRIBUTION

As the wave packet propagates in the field of the flow velocity gradient, one more intriguing property of the pulsation, i.e., the variation of the ratio of its components, becomes apparent. In the flow core, where the velocity gradient is negligible and the trajectories may therefore be considered rectilinear, set (6) transforms into a set of algebraic equations. The linearity of the trajectories means that randomly arranged incident and reflected waves superpose at any point of the channel. The only criterion for estimating the contribution of a wave to the disturbance is the attenuation of the wave, which depends on the distance l to the point of observation along the corresponding ray.

The complex velocity profile in the boundary layer greatly affects the wave motion. In this case, too, waves superpose to form the space–time pulsation field. However, here waves with curved trajectories interact with each other, so that their contribution to the disturbance depends on a number of factors in a complicated manner. In particular, one must know whether the point of disturbance origination is up- or downstream from the point of observation, what the distance between these two points is, and what the distance is between the point of observation and the wall.

Consider the section $y = 0$ in a plane-parallel channel. Through the point $z = z_*$, there passes an infinite number of trajectories with different initial values of xp. For each z_*, there exists a limiting value of x_*

beyond which ascending trajectories do not pass through this point z_*. The limiting trajectory has a limiting initial slope ϑ_*, which depends on M_0 and m. However, the point z_* is also crossed by waves following descending (not only ascending) trajectories. These waves also take part in the interaction and formation of the space–time pulsation field. Their path along these trajectories increases, and their role in disturbing hydrodynamic parameters becomes dependent on the position of the point of observation relative to the wall. It is essential that a change in the ray slope ϑ modifies the relationship between the contributions of the disturbances to the pulsation components.

Let us demonstrate this with a flow for which $M_0 = 0.05$ and $m = 5$ by contrasting the superposition of disturbances at the points $\{x, z\} = \{0, 0.5\}$ and $(x, z) = \{0, 0.1\}$. The trajectories of rays are constructed in Fig. 7. For simplicity, only waves originating at the lower wall are shown. Taking into account reflected walls does not change the chain of argument.

As follows from Fig. 7, the inclusion of the velocity profile $M(z)$ breaks the symmetry of trajectories about the point of observation $x = 0$. Disturbances downstream of this point on the right of the line $x = 0$ have initial coordinates $xp > 0$. Let us designate them as xp^+. Disturbances upstream of the point of observation are designated as xp^-. It is seen that disturbances from the domain xp^+ enhance the component u_z of the pulsation compared with the case of symmetric trajectories, while disturbances from the domain xp^- increase the component u_x.

Wave elements (rays) with the initial orientation covering the entire interval $\vartheta_0 \in [0, -90°]$ come to the point of observation from the domain xp^+. For rays coming to this point from the domain xp^-, the initial orientation is bounded from above by a value close to ϑ_*, $\vartheta_0 \in [0, \approx\vartheta_*]$. However, the extent of the domain xp^- is much greater than that of the domain xp^+. Accordingly, when disturbances **rp** are uniformly distributed over the wall and the flow velocity profile is taken into account, the pulsation vector **u** is oriented largely in the direction of the flow velocity **U**.

This tendency was discovered by Reichardt [5] in channel experiments and Klebanoff [6] in experiments with a plane-parallel plate [7, Chapter 18, §4]. Reichardt measured pulsation in the air flowing in a rectangular (1×0.244 m) channel. The Mach number was $M_0 = 0.003$ ($U_0 = 100$ cm/s), and the Reynolds number exceeded 10^4. The Reichardt's distributions of the rms longitudinal, $\sqrt{\bar{u}^2}$, and transverse, $\sqrt{\bar{w}^2}$, pulsations normalized to the maximal longitudinal pulsation $(\sqrt{\bar{u}^2})_{max}$ are shown in Fig. 8a.

The pulsation transverse component depends on z only weakly, while the longitudinal component peaks
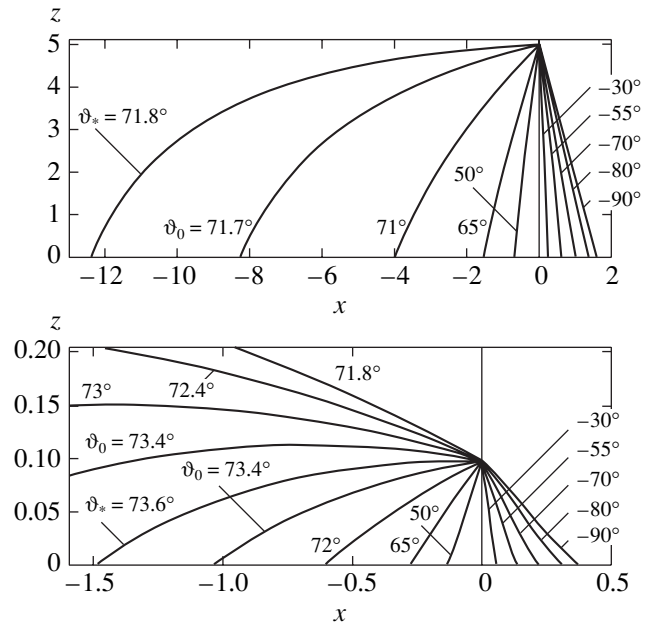

Fig. 7. Trajectories of rays approaching the points of observation $\{x, z\} = \{0, 0.5\}$ and $\{x, z\} = \{0, 0.1\}$.
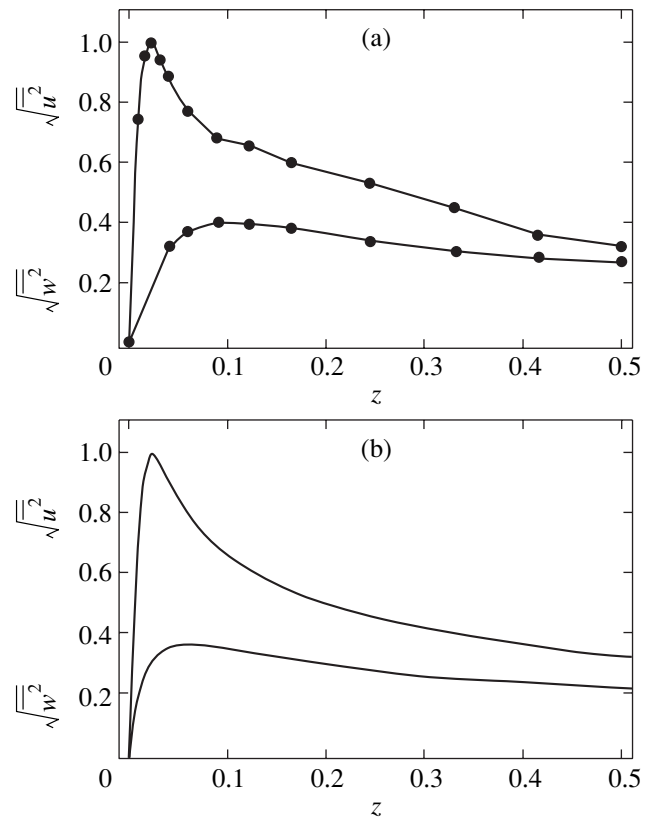

Fig. 8. Variation of the longitudinal, $\sqrt{\bar{u}^2}$, and transverse, $\sqrt{\bar{w}^2}$, velocity pulsations in the channel. (a) Reichardt's experiment and (b) simplified computational scheme.

sharply at a distance $\delta \approx 0.03$ from the wall. This finding, which has not yet been given an adequate explanation, can be readily treated in terms of the wave mechanism of turbulence. Moreover, this mechanism makes it possible to describe pulsation components on a quantitative basis.

Using relationship (1b), we write the velocity pulsation components as

$$u \sim f\left(\frac{l_n - l}{a}\right)\sin(\vartheta)/l,$$

$$w \sim f\left(\frac{l_n - l}{a}\right)\cos(\vartheta)/l. \tag{11}$$

Here, l is the distance traveled by the wave, l_n is the distance to the point of observation, ϑ is the slope of the ray at this point, and $a = \delta$ is the characteristic size of the primary disturbance. These parameters can be evaluated by solving Eqs. (6) and (1b). However, we will derive relationships for the components using a simplified computational scheme to demonstrate more pictorially how our method works. First, let us eliminate reflected waves from consideration and approximate the arc connecting the points $\{xp - x, z = 0\}$ and $\{xp - x, z\}$ by a straight line. The choice of the coordinate x of the point of observation is arbitrary, and the ordinate z varies within the half-height of the channel.

It is assumed that the distribution of xp along the x axis is uniform and the position of the primary disturbance is within the interval $xp \in [-10, 10]$. Outside this interval, the velocity pulsation decreases by more than $(xp/\delta) = 330$ times; so, it may be disregarded. Then, the z dependence of the rms pulsations can be recast by expressing the quantities entering into (11) through xp and z:

$$\sqrt{\bar{u}^2} \sim \sqrt{2\int_0^{10}\left(\frac{(1+\varepsilon)xp}{xp^2 + (z+a)^2}\right)^2 d(xp)},$$

$$\sqrt{\bar{w}^2} \sim \sqrt{2\int_0^{10}\left(\frac{(1-\varepsilon)xp}{xp^2 + (z+a)^2}\right)^2 d(xp)}. \tag{12}$$

Here, the coefficient ε compensates for changes in the contributions to the pulsation when the trajectories are bent by the flow velocity gradient. The parameter a in the denominator takes into account the fact that expression (1) is valid outside the boundary layer, so that one must put $z > a$ when integrating Eqs. (6). In constructing the graphs, we took into consideration that, within the primary disturbance, the field velocity drops from its maximum at $z = a$ to zero at $z = l = 0$. Such a pulsa-

tion distribution persists when disturbances making up the boundary layer superpose.

The pulsation components found by formulas (12) are shown in Fig. 8b (ε was taken to be 0.2). Comparing Figs. 8a and 8b, we conclude that the model and measured dependences are in good agreement. Note that $\varepsilon = 0.2$ is the result of a rather crude approximation. This value can be refined by using the experimentally found ratio of variously oriented pulsations. However, since ε depends on xp and z, the use of (12), instead of solving Eqs. (1) and (6), to obtain more exact values of the pulsation components is inappropriate.

CONCLUSIONS

To validate the mechanism discussed in this paper, let us estimate the parameters of an acoustic wave that can generate appropriate turbulent pulsations. In the Reichardt's measurements, the air flow had a velocity $U = 100$ cm/s and the velocity pulsation $u \approx 5$ cm/s. In an acoustic wave, the velocity and pressure pulsations are related through the relationship $u/c = p/p_0$ [4]; hence, $p/p_0 \sim 10^{-4}$. Then, for air under atmospheric pressure ($p_0 \sim 10^5$ Pa), we find $p \approx 10$ Pa or 0.1 mm Hg. The values $u \approx 5$ cm/s and $p \approx 10$ Pa found experimentally are by no means exotic. The human voice contains pulsations of such an amplitude.

Our concept of turbulent pulsation is based on acoustic wave superposition. As there are no special limitations involved, it may be applied to study pulsations in various media, for example, in low-temperature plasmas, where ion–sound waves are present.

REFERENCES

1. L. N. Pyatnitsky, Zh. Éksp. Teor. Fiz. **113**, 191 (1998) [JETP **86**, 107 (1998)].

2. L. N. Pyatnitsky, Zh. Éksp. Teor. Fiz. **119**, 665 (2001) [JETP **92**, 576 (2001)].

3. L. N. Pyatnitsky, Pis'ma Zh. Tekh. Fiz. **29** (9), 9 (2003) [Tech. Phys. Lett. **29**, 357 (2003)].

4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).

5. H. Reichardt, Messungen turbulenter Schwankungen Naturwissenschaften, 404 (1938).

6. P. S. Klebanoff, *Characteristics of Turbulence in Boundary Layer with Zero Pressure Gradient*, NASA Rep. 1247 (1955).

7. H. Schlichting, *Boundary Layer Theory* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1974).

Translated by V. Isaakyan

============================ **SOLIDS** ============================

Dynamic Properties of Ni, Cu, and Fe in the Condensed State: The Molecular Dynamics Method

A. G. Chirkov, A. G. Ponomarev, and V. G. Chudinov[†]

Udmurt State University, Izhevsk, 426037 Russia
e-mail: xps@fti.udm.ru
Received June 17, 2003

Abstract—The distribution function for the density of vibrational states and the velocity distribution function in the crystalline, liquid, and amorphous states of Ni, Cu, and Fe are studied by the molecular dynamics method. In the crystalline and amorphous states, the dynamic properties are qualitatively the same, while in the liquid state, additional low-energy excitations appear. These excitations may be treated as low-frequency resonant modes that arise because of the significant contribution of nonlinearity to the interaction potentials. In all the three states, the velocity distribution functions are found to be Maxwellian; that is, the systems are ergodic, although nonlinearity-related contributions to the interaction potentials are high and the liquid and amorphous states are disordered. © 2004 MAIK "Nauka/Interperiodica".

INTRODUCTION

Elucidation of amorphization mechanisms on the atomic level is a topical problem today. Amorphization is known to be associated with the formation and evolution of defects (vacancies, dendrites, or clusters) even in the liquid state, i.e., with the liquid state dynamics. Therefore, studying the liquid state of a system that is amorphous in the solid state is of crucial importance for understanding the process of amorphization. Dynamic properties of condensed media define their fundamental characteristics, such as diffusion, heat capacity, heat conduction, entropy, defect formation kinetics, scattering of quasi-particles by each other and by defects, etc. In the thermodynamic limit, these characteristics are completely described by the distribution function for the density of atomic vibrational states (DAVS) and the atomic velocity (or energy) distribution function (AVDF). In our opinion, it is appropriate to study the problem of amorphization in terms of nonlinear system dynamics, since the above processes are related to phase transitions, which are nonlinear effects. The liquid state is possible only when the elastic constant is equal to zero in the presence of nonlinearity: $d^2F/dV^2 = dp/dV = 0$ (where F is the free energy, V is the volume, and p is the pressure). This means (with no regard for the entropy contribution) that the interaction potential of an atom with all other atoms in the medium (on-side potential) has a flat or two-well bottom; that is, the atom does not have a fixed position. Media with harmonic interaction cannot be in the liquid state, because $d^2F/dV^2 = k$ (elastic constant). In this case, atoms sublimate from the free surface [1].

It has been shown [2, 3] that excitations of two types may arise in nonlinear systems, high-frequency local excitations (at a frequency above the cutoff frequency of a crystal) and low-frequency resonant excitations, depending on the type of nonlinearity (hard or soft). The former excitations are today called discrete breathers, which have been the subject of extensive research. They can be visualized as storages of the kinetic energy of a system's particles or "hot" areas with an elevated concentration of the kinetic energy. It is anticipated that such "defects" appear in amorphous materials and are responsible for long-term relaxation properties [4]. By using the molecular dynamics method, it was shown [5, 6] that such excitations may be present in high-temperature superconductors. As for low-frequency resonant modes, it was suggested [7] that they may diffuse, giving rise to the linear term in the expression for heat capacity at low temperatures in glasses. In other words, it is assumed that nonlinear excitations in nonlinear disordered systems may cover different frequency ranges and be responsible for the fundamental properties of these systems.

EXPERIMENTAL

Simulation was performed with a program package [8] in which emphasis is on the optimization of the time characteristics of the algorithm. A grain (crystallite) contains from 1000 to 2000 atoms, and cyclic boundary conditions are imposed. The system is brought to equilibrium at $T = 0$ K. The temperature is set by assigning velocities equal in magnitude but random in direction to the atoms. Then, the system is brought to equilibrium in a real time of $\sim 10^{-11}$ s. Within this time, the equilibrium DAVS and AVDF set in. The DAVS was calculated through the autocorrelation function [9]

[†] Deceased.

$G(w)$, arb. units
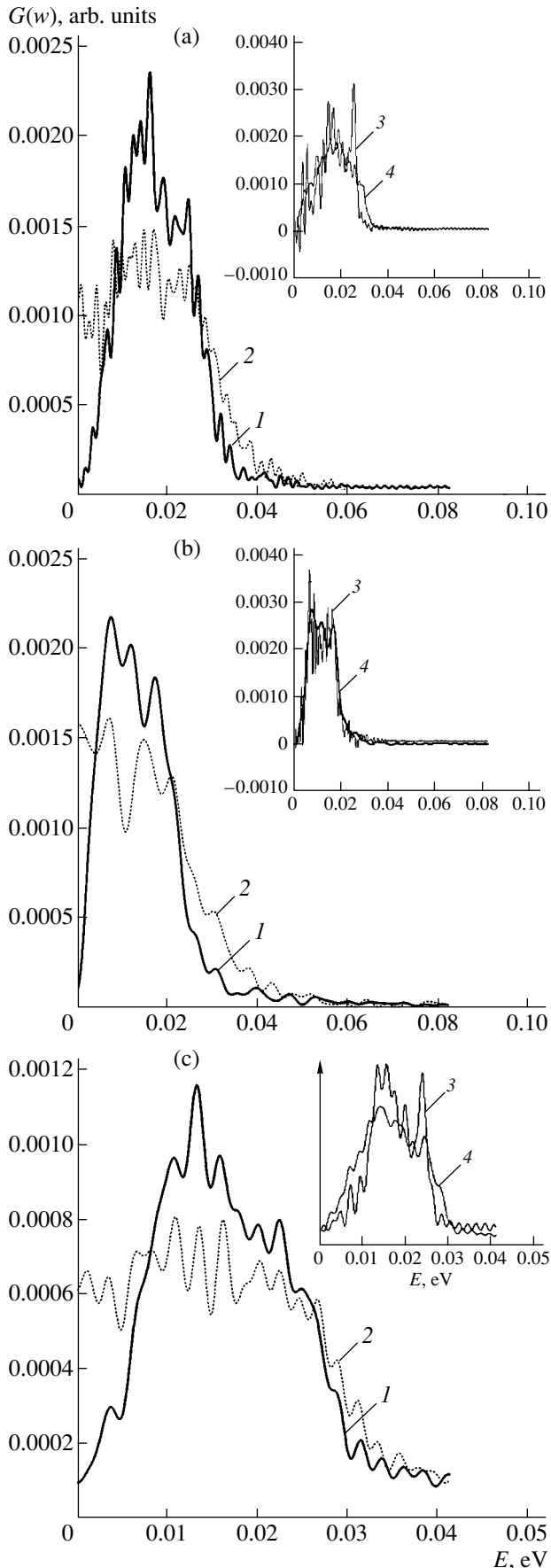
$$G(\omega) = \int_{0}^{\infty} \gamma(\tau) \exp(-i\omega\tau) d\tau, \qquad (1)$$

where

$$\gamma(\tau) = \sum_{i} \langle v_i(\tau) v_i(0) \rangle / v_i^2(0)$$

is the autocorrelation function, $v_i(\tau)$ is the velocity vector magnitude of an ith atom at a time τ, and $\langle \ldots \rangle$ means averaging over different time intervals.

Now the system is heated stepwise to the liquid state, further heated to a desired temperature, and then cooled stepwise to $T = 300$ K at a rate of $\sim 10^{-12}$ K/s, thus becoming amorphous.

The pair potential technique makes it possible to calculate the static and dynamic properties. At present, the molecular dynamics method employs model pair potentials adjusted to particular experimental data. We used pair potentials obtained in terms of the Heine–Aborenkov–Animalu pseudopotential approach [10]. In our opinion, they describe the entire set of structural and kinetic properties most adequately. As adjustable parameters, X-ray terms, electron density, and Show screening function were applied. In this case, the calculated basic characteristics of s, p, and d metals (lattice constants, elastic constants, phonon spectra, phase transition temperatures, energies of formation and migration of defects, etc.) differ from those found experimentally by 10–20%.

RESULTS AND DISCUSSION

Figure 1 shows the calculated DAVS for the crystalline, amorphous, and liquid states of Ni, Cu, and Fe. The drastic difference between the three DAVS curves is noteworthy. In the liquid state, the diffusion coefficient for Cu and Ni equals 5×10^{-5} cm²/s; for Fe, 10^{-5} cm²/s. The melting points calculated are $T_m = 1250$ K for Ni and 1500 K for Cu. The temperature at which the bcc phase of Fe becomes unstable is 1050 K. At the melting point, the low-energy part of the DAVS curve starts increasing with temperature; that is, additional low-frequency states appear in the system. Such a pattern is typical of all the metals. The occurrence of low-frequency states as low-frequency resonant modes was predicted within the framework of the anharmonic approach (see above). Their identification in our case is a challenge. It appears that, as the temperature grows, the conditions of high nonlinearity are established in the system and, once a critical point (the melting point) is reached, nonlinear excitations, i.e., nonlinear resonant modes, arise. In a homogeneous medium, these modes are not localized and dynamically appear at any

Fig. 1. Density of atomic vibrational states: (*1*) $T = T_m$, (*2*) $T = T_m + 100$ K, and (*3*) $T = 300$ K (crystalline state); (*4*) $T = 300$ K (quenched state). (a) Ni, (b) Cu, and (c) Fe.

lattice site. According to [2, 7], the energy accumulated in them is large enough for the atomic oscillation amplitude to be as high as the interatomic spacing. In this case, the atoms inevitably jump the potential barrier and the diffusion rate rises dramatically. Thus, one may say that the liquid state of a system is characterized by the excitation of nonlinear resonant modes. In linear systems, such an effect is obviously impossible. Nonlinear resonant modes appear if there exists a local critical volume (this explains the increase in the total liquid volume). Upon amorphization, the temperature conditions for nonlinear resonant mode excitation do not hold. Yet since the total volume of amorphous systems exceeds that of crystals, the former have voids, which provide local critical volumes for nonlinear resonant mode excitation. However, the number of these critical volumes is not large and they are uniformly distributed only in rarefied areas of, rather than throughout, the system. That is why the diffusion coefficients in amorphous systems are much lower than in liquids but higher than in crystals. High-frequency excitations in amorphous systems were not detected.

Figure 2 demonstrates the AVDFs for the crystalline, liquid, and amorphous state of Ni. They, as well as the AVDFs for Fe and Cu, are Maxwellian, which is surprising and leads one to nontrivial inferences. The fact is that amorphous materials are considered nonequilibrium (nonergodic). Whether a system is ergodic or not is important for theoretical analysis. Specifically, Gibbs classical statistical mechanics applies only to ergodic systems, for which the coefficient of correlation equals zero [11]. In general, a system is considered ergodic if the autocorrelation function is other than zero. Such systems can be defined as systems with memory. This concept is of special importance as applied to amorphous systems, because it is theorized that they inherit the properties of the liquid.

The Maxwell distribution is rigorously derived only under the assumption that systems or particles for which the coefficient of correlation equals zero are statistically independent to the maximum possible extent [12, 13]. Therefore, this distribution is also valid for ergodic systems. For such systems, bringing to equilibrium is described by the Boltzmann kinetic equation, the equilibrium solution to which is the Maxwell distribution. However, the applicability of the Boltzmann equation is limited by the short-range character of interaction forces ($\sim A/r^n$, where $n > 4$, for repulsion of atoms or for atomic interaction by the hard-sphere law [14]) and by introducing physically infinitesimal time and spatial domains in order to loose correlations [15]. Boltzmann completely ignored the correlations (the hypothesis for the number of collisions). In condensed media, the problem is reduced to considering quasi-particle gas. However, quasi-particle gas meets the above conditions only in a harmonic or quasi-harmonic approximation. In nonlinear systems, ergodicity is, in general, not obvious, since the assumptions underlying Gibbs statistical mechanics have to be justified for each
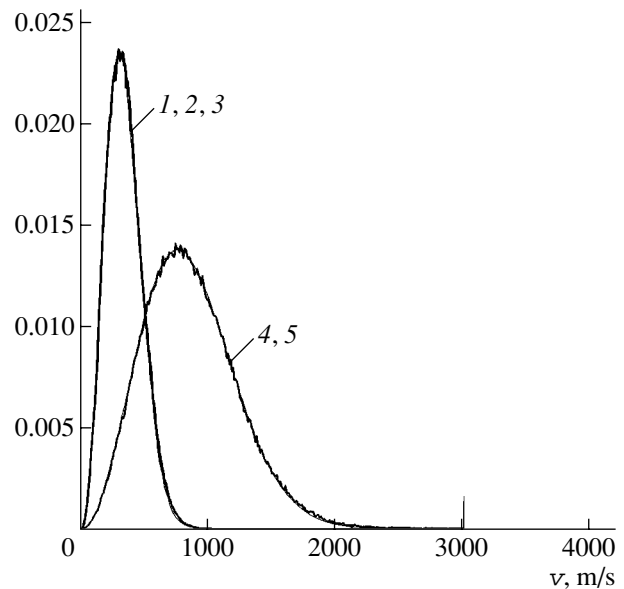

Fig. 2. Atomic velocity distribution function: (*1*) $T = 300$ K, Maxwellian; (*2*) $T = 300$ K, crystal; (*3*) $T = 300$ K, amorphous state; (*4*) $T = 2000$ K, Maxwellian; and (*5*) $T = 2000$ K, liquid.

specific problem. Although the Gibbs canonical distribution is merely postulated in practice, it turns out to be valid in most cases [16]. It was noted [17] that anharmonicities due to the violation of statistical independence may give rise to a non-Maxwellian particle velocity distribution and disturb the Gibbs canonical distribution. In solids, the distribution is, as a rule, Maxwellian; however, in nonlinear systems, it may not be so, as exemplified by LaSrCuO and YBaCuO high-temperature superconductors [17, 18].

Our results suggest that the systems investigated are ergodic. This is a very important conclusion, since classical thermodynamics, which is valid for ergodic systems, may be applied to these media in a dynamic approximation.

CONCLUSIONS

(1) Dynamically, the liquid state of a condensed system is characterized by the excitation of nonlinear low-frequency resonant modes. They cause the atoms to jump and, accordingly, accelerate diffusion in the liquid state.

(2) High-frequency excitations cannot be considered as a fundamental property of amorphous systems and liquids.

(3) The liquid and amorphous states are ergodic in a dynamic approximation, so that their behavior may be treated in terms of the classical thermodynamic approach.

ACKNOWLEDGMENTS

REFERENCES

1. Yu. K. Kovneristyĭ, É. K. Osipov, and E. A. Trifonov, *Physicochemical Foundations for Amorphous Metallic Alloys Production* (Nauka, Moscow, 1983).

2. A. J. Sievers and S. Takeno, Phys. Rev. Lett. **61**, 970 (1988).

3. A. M. Kosevich and A. S. Kovalev, Zh. Éksp. Teor. Fiz. **67**, 1793 (1974) [Sov. Phys. JETP **40**, 891 (1975)].

4. G. P. Tsironis and S. Aubry, Phys. Rev. Lett. **77**, 5225 (1996).

5. V. G. Chudinov, A. G. Chirkov, E. B. Dolgusheva, and V. M. Dyadin, Sverkhprovodimost **6**, 204 (1993).

6. A. G. Chirkov and V. G. Chudinov, Zh. Tekh. Fiz. **71** (1), 36 (2001) [Tech. Phys. **46**, 34 (2001)].

7. A. J. Sievers and S. Takeno, Phys. Rev. B **39**, 3374 (1989).

8. V. M. Dyadin, V. G. Chudinov, I. L. Gondyreva, and E. A. Bychkov, Available from VINITI No. 1537-B-91 (1991).

9. J. M. Dickey and A. Paskin, Phys. Rev. **188**, 1407 (1969).

10. *Particle Interaction in Liquid Metals*, Ed. by O. A. Esin (Nauka, Moscow, 1979).

11. V. I. Tikhonov, *Statistical Radio Engineering* (Sov. Radio, Moscow, 1966).

12. L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Nauka, Moscow, 1976; Pergamon, Oxford, 1980).

13. Yu. L. Klimontovich, *Statistical Physics* (Nauka, Moscow, 1982; Harwood Academic, New York, 1986).

14. *Nonequilibrium Phenomena: The Boltzmann Equation*, Ed. by L. J. Libovitz and E. U. Montroll (North-Holland, Amsterdam, 1983; Mir, Moscow, 1986).

15. N. N. Bogoliubov, *Problems of the Dynamic Theory in Statistical Physics* (GITTL, Moscow, 1946).

16. V. V. Kozlov, Dokl. Akad. Nauk **370**, 325 (2000).

17. V. G. Chudinov, A. G. Chirkov, and R. R. Nurgayanov, Fiz. Nizk. Temp. **24**, 13 (1998) [Low Temp. Phys. **24**, 9 (1998)].

18. E. B. Dolgusheva, V. G. Chudinov, and A. G. Chirkov, Fiz. Tverd. Tela (St. Petersburg) **41**, 1729 (1999) [Phys. Solid State **41**, 1585 (1999)].

Translated by V. Isaakyan

== **SOLIDS** ==

Thermodynamic Grounds for the Invar and Ellinvar Effects in Ferromagnets

V. Yu. Bodryakov and A. A. Povzner

Ural State Technical University (UPI), Yekaterinburg, 620002 Russia
e-mail: povz@kf.ustu.ru
Received July 10, 2003

Abstract—In terms of the standard thermodynamic approach (the Debye–Grüneisen model of solid and the Landau theory of second-order phase transitions), approximate conditions providing the constancy of the volumetric thermal expansion coefficient (the Invar effect) and of the bulk modulus (the Ellinvar effect) of a ferromagnet are established. Conditions under which a ferromagnet may exhibit Invar and Ellinvar properties simultaneously are found. Interaction between the magnetic, phonon, and electron subsystems of a ferromagnet is shown to be a crucial factor in the occurrence of the Invar and Ellinvar effects. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The Invar and Ellinvar effects in ferromagnets have drawn the attention of researchers for many years because of their scientific and applied significance (see, e.g., [1–8] and references therein). Although these effects are widely used in practice and are treated in terms of various models, the conditions under which they take place still remain unclear. As far as we know, even a consistent thermodynamic consideration of these effects is lacking. As a result, thermodynamic reasons for the occurrence of these effects and the role of ferromagnetic ordering remain a mystery. Unfortunately, no consideration is given to the Invar and Ellinvar problems in the classical course of theoretical physics (neither in the first [9] nor in the latest [10] editions).

The aim of this work is to show that the Invar and Ellinvar effects are a natural consequence of ferromagnetic ordering and occur given a specific relationship between thermodynamic parameters. In doing so, we do not make any specific assumptions regarding the "structure" of a ferromagnet and do not go beyond the scope of the standard thermodynamic concepts (the Debye–Grüneisen model of solid and the Landau theory of second-order phase transitions.

1. THEORY

In this section, we will obtain (in terms of simple concepts of solid ferromagnetic metal) thermodynamically correct expressions for the first and second derivatives of the thermodynamic potential (free energy) of a ferromagnetic metal. It will be formally assumed that the Debye temperature θ depends on the temperature and volume (pressure) but not on the magnetization M and also that the M dependence of θ can be reduced to a relevant temperature dependence through the equa-

tion of magnetic state (EMS). In addition, the temperature dependences $\theta = \theta(T)$ are assumed to be different in the ferromagnetic (FM) and paramagnetic (PM) ranges of a magnet. The general dependence of the characteristic temperature θ on the temperature and volume of a paramagnet was discussed in detail elsewhere [11–17] and is not considered in this work. In a number of cases, the temperature dependence of the Debye temperature may be disregarded within a limited temperature interval for both the magnetically ordered and paramagnetic ranges of a ferromagnet.

(i) First thermodynamic derivatives of the free energy and thermodynamic potential of a ferromagnet. We proceed from the conventional differential representation of the free molar energy F (as a function of temperature, molar volume, and magnetic field) and the thermodynamic potential Φ (as a function of temperature, pressure, and magnetic field) [9, 10]:

$$dF(T, V, B) = -SdT - PdV - MdH, \qquad (1)$$

$$d\Phi(T, P, B) = -SdT + VdP - MdH, \qquad (2)$$

where S is the molar entropy, M is the molar magnetization, and H is the magnetic field normalized appropriately.

The conventional integral additive representation of the free energy and thermodynamic potential per mole are given by

$$F = F_{\text{para}} + F_{\text{m}}; \qquad (3)$$

$$\Phi = \Phi_{\text{para}} + \Phi_{\text{m}}. \qquad (4)$$

Here, $F_{\text{para}} = F_0 + F_1 + F_e$ and $\Phi_{\text{para}} = \Phi_0 + \Phi_1 + \Phi_e$ are the "paramagnetic cores" of the free energy and thermodynamic potential, respectively; $F_0 = F_0(V)$ and $\Phi_0 = \Phi_0(P)$ are the "constant" (temperature and magnetization independent) contributions to the free energy

and thermodynamic potential, respectively; $F_l(T, \theta)$ and $\Phi_l(T, \theta)$ are the lattice (phonon) parts of the free energy and thermodynamic potential, respectively; and $F_e(T)$ and $\Phi_e(T)$ are the electron contributions. All the contributions were repeatedly discussed previously (see, e.g., [16]).

With allowance for the Zeeman contributions in terms of the Landau theory of second-order phase transitions [16], we can write for the magnetic components of the free molar energy and thermodynamic potential

$$F_m = \frac{1}{2}\alpha(T, V)M^2 + \frac{1}{4}\beta(T, V)M^4 - MH, \qquad (5)$$

$$\Phi_m = \frac{1}{2}\alpha(T, P)M^2 + \frac{1}{4}\beta(T, P)M^4 - MH. \qquad (6)$$

In these expressions, α and β are thermodynamic coefficients that generally depend on the temperature and volume (for free energy) or (pressure (for thermodynamic potential). In the Landau approximation it is assumed that $\alpha = a(T - T_C)$, $a > 0$, and $\beta > 0$ and that the Curie temperature T_C, as well as the thermodynamic coefficients, depends on the volume (pressure). Note that the idea of the constancy of the thermodynamic coefficients a and β is contrary to fact, as was noted in [2]; therefore, their temperature dependence should be taken into account. Below, we assume that this dependence is arbitrary (other than that suggested by Landau) and obtain results that are consistent with the Landau approximation.

The equilibrium magnetization value, which is the order parameter in the case of a ferromagnet, is found by minimizing thermodynamic potential (6) when the temperature, pressure, and magnetic field are constant:

$$\left(\frac{\partial \Phi}{\partial M}\right)_{TPH} = 0 = \alpha M + \beta M^3 - H. \qquad (7)$$

Equation (7) is called the equation of magnetic state or the Belov–Arrott equation. Solving it analytically or numerically, one finds the molar magnetization as a function of temperature and magnetic field.

In the absence of the magnetic field ($H = 0$) in the magnetically ordered range ($T \leq T_C$), the spontaneous magnetization is conventionally given by

$$M_s^2 = -\frac{\alpha}{\beta}; \qquad (8)$$

above T_C, the spontaneous magnetization $M_s = 0$. In the Landau approximation, the spontaneous magnetization takes the form

$$M_s^2 = -\frac{a}{\beta}t, \qquad (9)$$

where $t = T - T_C$ is the deviation from the Curie temperature.

From the above expressions for free energy and thermodynamic potential, we can find thermodynamically correct (in terms of the Landau theory) expressions for their first thermodynamic derivatives, namely, for the molar entropy S, molar volume V, and pressure P:

$$S = -\left(\frac{\partial \Phi}{\partial T}\right)_{PH} = -\left(\frac{\partial \Phi_{para}}{\partial T}\right)_{PH} - \left(\frac{\partial \Phi_m}{\partial T}\right)_{PH} \qquad (10)$$
$$= S_{para} + S_m;$$

$$V = \left(\frac{\partial \Phi}{\partial P}\right)_{TH} = \left(\frac{\partial \Phi_{para}}{\partial P}\right)_{TH} + \left(\frac{\partial \Phi_m}{\partial P}\right)_{TH} \qquad (11)$$
$$= V_{para} + V_m;$$

$$P = -\left(\frac{\partial F}{\partial V}\right)_{TB} = -\left(\frac{\partial F_{para}}{\partial V}\right)_{TH} - \left(\frac{\partial F_m}{\partial V}\right)_{TH} \qquad (12)$$
$$= P_{para} + P_m.$$

Here, S_{para}, V_{para}, and P_{para} are the paramagnetic components of the molar entropy, molar volume, and pressure, respectively, which involve, as well as the free energy and thermodynamic potential, the constant phonon and electron contributions.

These expressions were repeatedly analyzed previously (see, e.g., [16]). As an illustration, we give here the expression for the molar volume of the paramagnetic phase

$$V_{para} = V_0 + 3R\left[\frac{3}{8} + \frac{D(z)}{z}\right]\left(\frac{\partial \theta}{\partial P}\right)_T - \frac{1}{2}\left(\frac{\partial \zeta}{\partial P}\right)_T T^2, \quad (13)$$

where $z = \theta/T$ and ζ is the molar electronic heat.

In view of (7), the thermodynamically exact (in terms of the Landau theory) expressions for the magnetic components of the first thermodynamic derivatives of the free energy and thermodynamic potentials (S_m, V_m, and P_m) have the form

$$S_m = -\frac{1}{2}\left(\frac{\partial \alpha}{\partial T}\right)_{PH} M^2 - \frac{1}{4}\left(\frac{\partial \beta}{\partial T}\right)_{PH} M^4, \qquad (14)$$

$$V_m = \frac{1}{2}\left(\frac{\partial \alpha}{\partial P}\right)_{TH} M^2 + \frac{1}{4}\left(\frac{\partial \beta}{\partial P}\right)_{TH} M^4, \qquad (15)$$

$$P_m = -\frac{1}{2}\left(\frac{\partial \alpha}{\partial V}\right)_{TH} M^2 - \frac{1}{4}\left(\frac{\partial \beta}{\partial V}\right)_{TH} M^4. \qquad (16)$$

In the Landau approximation, these expressions become

$$S_m = -\frac{1}{2}aM^2, \qquad (17)$$

$$V_{\mathrm{m}} = -\frac{1}{2}a\left(\frac{\partial T_{\mathrm{C}}}{\partial P}\right)_{TH}M^2$$
$$+\frac{1}{2}\left(\frac{\partial a}{\partial P}\right)_{TH}M^2 t + \frac{1}{4}\left(\frac{\partial \beta}{\partial P}\right)_{TH}M^4, \tag{18}$$

$$P_{\mathrm{m}} = \frac{1}{2}a\left(\frac{\partial T_{\mathrm{C}}}{\partial V}\right)_{TH}M^2$$
$$+\frac{1}{2}\left(\frac{\partial a}{\partial V}\right)_{TH}M^2 t - \frac{1}{4}\left(\frac{\partial \beta}{\partial V}\right)_{TH}M^4. \tag{19}$$

An important and as yet uncovered (in the authors' opinion) issue is that, in the Landau theory, the magnetic component of the molar volume (as well as of the pressure) can be represented as a double series in even powers of the order parameter and deviation from the Curie temperature:

$$V_{\mathrm{m}} = \frac{1}{2}V_{10}M^2 + \frac{1}{2}V_{11}M^2 t + \frac{1}{4}V_{20}M^4, \tag{20}$$

where, as follows from (20), the coefficients of the series are

$$V_{10} = -a\left(\frac{\partial T_{\mathrm{C}}}{\partial P}\right)_{TH}, \quad V_{11} = \left(\frac{\partial a}{\partial P}\right)_{TH},$$

$$V_{20} = \left(\frac{\partial \beta}{\partial P}\right)_{TH}.$$

In the Landau approximation, the thermodynamic coefficient V_{ij}, which could be named the coefficients of magnetovolume interaction, are independent of temperature and magnetization.

In view of relationship (9), the terms of the series with the coefficients V_{11} and V_{20} in the absence of the field have a single (second) order of smallness; therefore, the expression for the magnetic component of the molar volume allows for two equivalent representations in the Landau approximation: either in even powers of the spontaneous magnetization,

$$V_{\mathrm{ms}} = \frac{1}{2}V_{10}M_{\mathrm{s}}^2 + \frac{1}{4}\left(V_{20} - \frac{\beta}{a}V_{11}\right)M_{\mathrm{s}}^4, \tag{21}$$

or in powers of deviation from the Curie temperature,

$$V_{\mathrm{ms}} = -\frac{1}{2}\frac{a}{\beta}V_{10}t + \frac{1}{4}\frac{a}{\beta}\left(\frac{a}{\beta}V_{20} - 2V_{11}\right)t^2. \tag{22}$$

From the above expressions, one can draw definite conclusions. Specifically, if the coefficient at t^2 in expression (22)

$$\frac{a}{\beta}V_{20} - 2V_{11} = a\left[\frac{1}{\beta}\left(\frac{\partial \beta}{\partial P}\right)_{T,\,H=0} - \frac{2}{a}\left(\frac{\partial a}{\partial P}\right)_{T,\,H=0}\right]$$

equals zero or, in other words, if the pressure dependence of the thermodynamic coefficients a and β is

such that the ratio β/a^2 is pressure independent in a first approximation, the spontaneous magnetostriction $\omega_{\mathrm{s}} = V_{\mathrm{ms}}/V$ will be a near-linear function of temperature in the ferromagnetic range. This property may be used, for example, in situations where it is necessary to provide a precision linear displacement of an object that is proportional to the temperature. If the condition $V_{10} > 0$ or $(\partial T_{\mathrm{C}}/\partial P)_{TH} < 0$ (which is the same) is also met, the magnetic part of the volumetric thermal expansion coefficient will be negative (see below). Such behavior is typical of Invar-like ferromagnets. These thermodynamic conditions may be set up by appropriately choosing the chemical composition of a ferromagnetic alloy.

(ii) Second thermodynamic derivatives of the free energy and thermodynamic potential of a ferromagnet. From the results obtained above and in view of (7), we will find the thermodynamically exact (in terms of the Landau theory) expressions for the second derivatives of the free energy and thermodynamic potential, namely, for the molar heat capacity

$$C = T(\partial S/\partial T)_{PH},$$

volumetric thermal expansion coefficient (VTEC)

$$o = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{PH},$$

and the bulk modulus

$$K = -V\left(\frac{\partial P}{\partial V}\right)_{TH}.$$

The thermodynamic quantities C, o, and K admit the additive representation

$$C = -T\left(\frac{\partial^2 \Phi}{\partial T^2}\right)_{PH} = -T\left(\frac{\partial^2 \Phi_{\mathrm{para}}}{\partial T^2}\right)_{PH} - T\left(\frac{\partial^2 \Phi_{\mathrm{m}}}{\partial T^2}\right)_{PH} \tag{23}$$
$$= C_{\mathrm{para}} + C_{\mathrm{m}},$$

$$o = \frac{1}{V}\left(\frac{\partial^2 \Phi}{\partial P\partial T}\right)_{H} = \frac{1}{V}\left(\frac{\partial^2 \Phi_{\mathrm{para}}}{\partial P\partial T}\right)_{H} + \frac{1}{V}\left(\frac{\partial^2 \Phi_{\mathrm{m}}}{\partial P\partial T}\right)_{H} \tag{24}$$
$$= o_{\mathrm{para}} + o_{\mathrm{m}},$$

$$K = V\left(\frac{\partial^2 F}{\partial V^2}\right)_{TH} = V\left(\frac{\partial^2 F_{\mathrm{para}}}{\partial V^2}\right)_{TH} + V\left(\frac{\partial^2 F_{\mathrm{m}}}{\partial V^2}\right)_{TH} \tag{25}$$
$$= K_{\mathrm{para}} + K_{\mathrm{m}}.$$

In (23)–(25), C_{para}, o_{para}, and K_{para} are the paramagnetic components of the molar heat capacity, VTEC, and bulk modulus, respectively, which involve the constant, phonon, and electronic contributions. Detailed analysis of the associated expressions was made elsewhere (see, e.g., [16]) and is omitted here. Here, we

need the expressions for o_{para} and K_{para}:

$$o_{\text{para}} = 3RC_{VR}(z)\left[1 - \frac{1}{z}\left(\frac{\partial\theta}{\partial T}\right)_{PH}\right]\frac{1}{\theta}\left(\frac{\partial\theta}{\partial P}\right)_{TH}$$

$$+ 3R\left[\frac{3}{8} + \frac{D(z)}{z}\right]\frac{\partial^2\theta}{\partial P\partial T} - \left(\frac{\partial\zeta}{\partial P}\right)_{TH}T, \tag{26}$$
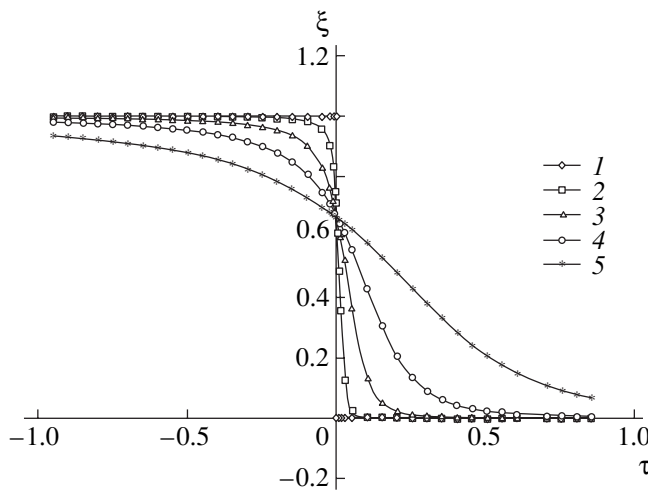
$$K_{\text{para}} = K_0 - 3RVT\left\{\left(\frac{1}{\theta}\frac{\partial\theta}{\partial V}\right)_{TH}^2 C_{VR}(z)\right.$$

$$\left. - \left[\frac{3}{8}z + D(z)\right]\frac{1}{\theta}\left(\frac{\partial^2\theta}{\partial V^2}\right)_{TH}\right\} - \frac{1}{2}V\left(\frac{\partial^2\zeta}{\partial V^2}\right)_{TH}T^2. \tag{27}$$

In this approximation and in view of Eq. (7) of magnetic state, we have for the magnetic components of the molar heat capacity, VTEC, and bulk modulus

$$C_{\text{m}} = -T\left[\frac{1}{2}\left(\frac{\partial^2\alpha}{\partial T^2}\right)_{PH}M^2 + \frac{1}{4}\left(\frac{\partial^2\beta}{\partial T^2}\right)_{PH}M^4\right]$$

$$+ \frac{\xi T}{2\beta}\left[\left(\frac{\partial\alpha}{\partial T}\right)_{PB} + \left(\frac{\partial\beta}{\partial T}\right)_{PB}M^2\right]^2, \tag{28}$$

$$o_{\text{m}} = \frac{1}{V}\left[\frac{1}{2}\left(\frac{\partial^2\alpha}{\partial P\partial T}\right)_H M^2 + \frac{1}{4}\left(\frac{\partial^2\beta}{\partial P\partial T}\right)_H M^4\right]$$

$$- \frac{\xi}{2\beta V}\left[\left(\frac{\partial\alpha}{\partial T}\right)_{PH} + \left(\frac{\partial\beta}{\partial T}\right)_{PH}M^2\right]\left[\left(\frac{\partial\alpha}{\partial P}\right)_{TH} + \left(\frac{\partial\beta}{\partial P}\right)_{TH}M^2\right], \tag{29}$$

$$K_{\text{m}} = V\left[\frac{1}{2}\left(\frac{\partial^2\alpha}{\partial V^2}\right)_{TH}M^2 + \frac{1}{4}\left(\frac{\partial^2\beta}{\partial V^2}\right)_{TH}M^4\right]$$

$$- \frac{\xi V}{2\beta}\left[\left(\frac{\partial\alpha}{\partial V}\right)_{TH} + \left(\frac{\partial\beta}{\partial V}\right)_{TH}M^2\right]^2. \tag{30}$$

Calculated dependences of the thermodynamic parameter ξ on the reduced temperature $\tau = T/T_C$ for a magnetic field $H =$ (1) 0, (2) 300, (3) 2500, (4) 10000, and (5) 40000 arb. units. The parameters $a = 10$ arb. units and $\beta = 1$ arb. unit.

In (28)–(30),

$$\xi = \frac{2\beta M^2}{\alpha + 3\beta M^2} \tag{31}$$

is the dimensionless thermodynamic parameter depending on the temperature and magnetization. The value of ξ is close to unity in the ferromagnetic range and to zero in the paramagnetic range (see figure). In the absence of the magnetic field, the exact values, as readily follows from (8) and (31), are $\xi = 1$ at $T < T_C$, $\xi = 2/3$ at $T = T_C$, and $\xi = 0$ at $T > T_C$. Such a step variation of ξ allows one to use it as the order parameter characterizing a ferromagnet.

In the Landau approximation, these complex expressions in the ferromagnetic range take the form

$$C_{\text{m}} = \frac{\xi a^2 T}{2\beta}; \tag{32}$$

$$o_{\text{m}} = \frac{\xi a^2}{2\beta V}\left(\frac{\partial T_C}{\partial P}\right)_{TH} + \frac{1}{2V}\left[\left(\frac{\partial a}{\partial P}\right)_{TH}\right.$$

$$\left. - \frac{\xi a}{\beta}\left(\frac{\partial\beta}{\partial P}\right)_{TH}\right]M^2 - \frac{\xi a}{2\beta V}\left(\frac{\partial a}{\partial P}\right)_{TH}t, \tag{33}$$

$$K_{\text{m}} = K_{00} + \frac{1}{2}K_{10}M^2 + K_{01}t + \frac{1}{4}K_{10}M^4$$

$$+ \frac{1}{2}K_{11}M^2 t + \frac{1}{2}K_{02}t^2. \tag{34}$$

The expansion of K_{m} in M^2 and t is the most tedious. Its coefficients are

$$K_{00} = -\frac{\xi a^2 T_C^2}{2\beta V}\gamma_C^2, \tag{35}$$

$$K_{10} = \frac{aT_C}{V}(-2\gamma_a\gamma_C + 2\gamma_\beta\gamma_C - \gamma_C^*), \tag{36}$$

$$K_{01} = \frac{\xi a^2 T_C}{\beta V}\gamma_a\gamma_C, \tag{37}$$

$$K_{20} = \frac{\beta}{V}(\gamma_\beta^* - 2\xi\gamma_\beta^2), \tag{38}$$

$$K_{11} = \frac{a}{V}(-2\gamma_a\gamma_\beta + \gamma_a^*), \tag{39}$$

$$K_{02} = \frac{\xi a^2}{\beta V}\gamma_a^2. \tag{40}$$

To make expressions (35)–(40) more compact, we designated the first and second dimensionless isothermal derivatives of the thermodynamic coefficients with respect to volume as

$$\gamma_i = \frac{V}{i}\left(\frac{\partial i}{\partial V}\right)_{TH} \quad \text{and} \quad \gamma_i^* = \frac{V}{i}\left(\frac{\partial i}{\partial V}\right)_{TH},$$

where $i = a, \beta, T_C$. For example,

$$\gamma_C = \frac{V}{T_C}\left(\frac{\partial T_C}{\partial V}\right)_{TH}.$$

Frequently, it is sufficient to consider the quantities γ_i and γ_i^* as temperature- and magnetization-independent parameters of a ferromagnet. It should be noted that, in the case of a solid, the thermodynamic derivatives with respect to pressure (e.g., expression (33)) may be easily transformed into the derivatives with respect to volume and vice versa (see, e.g., [16]). For example, the first derivatives of an arbitrary thermodynamic function f are related to each other as

$$\left(\frac{\partial f}{\partial P}\right)_{TH} = -\frac{f}{K}\gamma_f.$$

For our study, the molar heat capacity is of no significance, so we will analyze only the VTEC and bulk modulus of a ferromagnet. In the absence of the field ($H = 0$) (the case of most practical interest), the spontaneous magnetic part of the VTEC can be written (in view of the above results) either as

$$o_{ms} = \frac{\xi a^2}{2\beta V}\left(\frac{\partial T_C}{\partial P}\right)_{TH}$$
$$+ \frac{a}{2V}\left[\frac{2}{a}\left(\frac{\partial a}{\partial P}\right)_{TH} - \frac{1}{\beta}\left(\frac{\partial \beta}{\partial P}\right)_{TH}\right]M_S^2 \qquad (41)$$

or as

$$o_{ms} = \frac{\xi a^2}{2\beta V}\left(\frac{\partial T_C}{\partial P}\right)_{TH}$$
$$- \frac{a^2}{2V\beta}\left[\frac{2}{a}\left(\frac{\partial a}{\partial P}\right)_{TH} - \frac{1}{\beta}\left(\frac{\partial \beta}{\partial P}\right)_{TH}\right]t. \qquad (42)$$

Thus, the VTEC of a ferromagnet experiences a jump at the Curie temperature (the sign of the jump coincides with the sign of the derivative with respect to pressure $\left(\frac{\partial T_C}{\partial P}\right)_{TH}$ and then varies in direct proportion to the spontaneous magnetization squared or, which is equivalent in the Landau approximation, to the deviation from the Curie temperature $t = T - T_C$) as the temperature decreases in the ferromagnetic range. The constancy condition for the magnetic part of the VTEC in the ferromagnetic range is evident: it was given in the analysis of the expression for the spontaneous magnetic part of the molar volume.

As follows from expressions (9) and (34), the spontaneous contribution to the bulk modulus in the magnetically ordered range of a ferromagnet in the Landau approximation can be represented, like o_{ms}, in two

equivalent forms:

$$K_m = -\frac{a^2 T_C^2}{2\beta V}\gamma_C^2 - \frac{aT_C}{V}\left(2\gamma_a\gamma_C - \gamma_C\gamma_\beta + \frac{1}{2}\gamma_C^*\right)M_s^2$$
$$- \frac{1}{2}\frac{\beta}{V}\left((\gamma_a - \gamma_\beta)^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^*\right)M_s^4, \qquad (43)$$

or

$$K_m = -\frac{a^2 T_C^2}{2\beta V}\gamma_C^2 + \frac{a^2 T_C}{\beta V}\left(2\gamma_a\gamma_C - \gamma_C\gamma_\beta + \frac{1}{2}\gamma_C^*\right)t$$
$$- \frac{1}{2}\frac{a^2}{\beta V}\left((\gamma_a - \gamma_\beta)^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^*\right)t^2. \qquad (44)$$

It is easy to check that, in the absence of the field, the magnetic part of the bulk modulus of a ferromagnet steps down at the Curie point in going to the magnetically ordered range. The further behavior of the modulus depends on the sign of the thermodynamic parameters and the relationship between them: the modulus may rise, as in gadolinium [18–22], or decline, as in nickel [1, 2]. In addition, the bulk modulus may remain almost constant within a certain temperature interval, as is usually observed in Ellinvar-like alloys [1, 2]. In the ferromagnetic range, the magnetic component of the bulk modulus will be roughly constant if the two thermodynamic conditions

$$2\gamma_a\gamma_C - \gamma_C\gamma_\beta + \frac{1}{2}\gamma_C^* = 0,$$

$$(\gamma_a - \gamma_\beta)^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^* = 0$$

are met simultaneously.

It is of interest that not only the magnetic part of the bulk modulus but also that of the VTEC may remain constant. In terms of the γ parameters, the latter condition has the form $\gamma_\beta - 2\gamma_a = 0$. For such a ferromagnet, the equalities $\gamma_C^* = 0$ and $\gamma_a^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^* = 0$ must also hold.

2. THERMODYNAMIC CONDITIONS FOR THE OCCURRENCE OF THE INVAR AND ELLINVAR EFFECTS IN A HIGH-TEMPERATURE FERROMAGNET

To reliably determine the conditions for the Invar and Ellinvar effects in ferromagnets, one should take into account not only the variation of the VTEC and bulk modulus due to magnetic ordering but also their temperature variation associated with the phonon and electron subsystems of a ferromagnet. In the general case, the resultant expressions are very awkward. Therefore, by way of example, we will consider only the case that seems to be of greatest practical interest (see table), that is, the case of a "high-temperature" fer-

Debye, θ, and Curie, T_C, temperatures for typical metallic ferromagnets [22]

Ferromagnet	θ, K	T_C, K
Gd	182	293.4
Fe	477	1044
Co	460	1388
Ni	477	627.4

romagnet (for which $\theta \le T_C$). Analysis will be carried out for elevated temperatures: $\theta \le T \le T_C$. Certainly, the above thermodynamic results are also valid in other temperature ranges (within the applicability domain of the Landau theory), for which numerical calculation, rather than analytical consideration, is more appropriate.

Thus, for a high-temperature ferromagnet kept at elevated temperatures ($\theta \le T \le T_C$), the paramagnetic components of the VTEC and bulk modulus are approximated as

$$o_{\text{para}} \approx -\frac{3R}{VK}\gamma_\theta + \frac{\zeta}{VK}\gamma_\zeta T, \tag{45}$$

$$K_{\text{para}} \approx K_0 - \frac{3RT}{V}(\gamma_\theta^2 - \gamma_\theta^*) - \frac{1}{2}\frac{\zeta\gamma_\zeta^*}{V}T^2. \tag{46}$$

Here, the subscript i introduced for the γ_i parameters (see above) runs through θ and ζ. Eventually, the general expression for the VTEC of a ferromagnet under the conditions mentioned above appears as

$$o \approx -\frac{3R}{VK}\left[\frac{a^2 T_C}{6\beta R}(2\gamma_a - \gamma_\beta + \gamma_C) - \gamma_\theta\right] \\ + \frac{\zeta T}{VK}\left[\frac{a^2}{2\beta\zeta}(2\gamma_a - \gamma_\beta) + \gamma_\zeta\right]. \tag{47}$$

In the context of this work, the temperature dependence of the bulk modulus is easier to analyze by taking its derivative to eliminate a high constant contribution K_0 to the total value of K:

$$\left(\frac{\partial K}{\partial T}\right)_{PH} \approx \frac{3R}{V}\left[\gamma_\theta^* - \gamma_\theta^2 + \frac{a^2 T_C}{3R\beta}\left(2\gamma_a\gamma_C - \gamma_C\gamma_\beta\right. \right. \\ \left. + (\gamma_a - \gamma_\beta)^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^* + \frac{1}{2}\gamma_C^*\right)\right] \tag{48}$$

$$- \frac{\zeta T}{V}\left[\frac{a^2}{\beta\zeta}\left((\gamma_a - \gamma_\beta)^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^*\right) + \gamma_\zeta^*\right].$$

Thus, the thermodynamic conditions under which the Invar and Ellinvar effects arise in ferromagnets can be written by equating the bracketed dimensionless criteria in (47) and (48) to zero. A ferromagnet will have a

roughly constant VTEC if

$$\psi_{o1} = \frac{a^2}{2\beta\zeta}(2\gamma_a - \gamma_\beta) + \gamma_\zeta = 0. \tag{49}$$

The VTEC of a ferromagnet will be close to zero if, additionally,

$$\psi_{o0} = \frac{a^2 T_C}{6\beta R}(2\gamma_a - \gamma_\beta + \gamma_C) - \gamma_\theta = 0. \tag{50}$$

From expressions (49) and (50), it is seen that the criteria ψ_{o0} and ψ_{o1} may vanish simultaneously if the parameter γ_θ (which coincides up to sign with the Grüneisen parameter Γ_θ: $\Gamma_\theta = -\gamma_\theta$) is related to the magnetic and (to a lesser extent) electronic properties of a ferromagnet as

$$\gamma_\theta = \frac{\zeta\gamma_\zeta T_C}{3R} - \frac{a^2 T_C}{6\beta R}. \tag{51}$$

This relationship indicates that correctly taking into account interaction between the magnetic, phonon, and electron subsystems of a magnet is of crucial importance in analyzing the conditions for the Invar effect.

Similarly, the derivative of the bulk modulus of a ferromagnetic will be roughly equal to zero if

$$\psi_{K0} = \gamma_\theta^* - \gamma_\theta^2 + \frac{a^2 T_C}{3R\beta}\left(2\gamma_a\gamma_C - \gamma_C\gamma_\beta \right. \\ \left. + (\gamma_a - \gamma_\beta)^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^* + \frac{1}{2}\gamma_C^*\right) = 0, \tag{52}$$

$$\psi_K = \frac{a^2}{\beta\zeta}\left((\gamma_a - \gamma_\beta)^2 + \gamma_a^* - \frac{1}{2}\gamma_\beta^*\right) + \gamma_\zeta^* = 0. \tag{53}$$

From (52) and (53), it follows that the criteria ψ_{K0} and ψ_{K1} vanish simultaneously only if

$$\gamma_\theta^2 - \gamma_\theta^* = \frac{a^2 T_C}{6R\beta}(4\gamma_a\gamma_C - 2\gamma_C\gamma_\beta + \gamma_C^*) - \frac{\zeta\gamma_\zeta^* T_C}{3R}. \tag{54}$$

Like the Invar effect, the Ellinvar effect takes place only if the magnetic, phonon, and electron subsystems of a ferromagnet meet certain relationships.

Note also that both effects can be observed in ferromagnets if thermodynamic conditions (49), (50), (52), and (53) are satisfied simultaneously.

Thus, as follows from our results, new ferromagnets with the Invar or Ellinvar properties can be produced by varying the chemical composition of the materials. For example, iron–nickel alloys, such as conventional Invar (36 wt % Ni in the iron matrix), may be alloyed with certain impurities in order to find component ratios providing the thermodynamic conditions for these effects.

CONCLUSIONS

Thus, using the consistent thermodynamic approach, we found thermodynamic criteria for the Invar (the temperature independence of the thermal expansion coefficient) and Ellinvar (the temperature independence of the bulk modulus) effects. The conditions under which both effects may occur simultaneously are established. Analytical results indicate that interaction between the magnetic, phonon, and electron subsystems of a ferromagnet is responsible for these effects in ferromagnets. This interaction must be taken into account upon elaborating microscopic models of the effects.

REFERENCES

1. K. P. Belov, *Elastic, Thermal, and Electrical Phenomena in Ferromagnetic Metals* (GITTL, Moscow, 1951).

2. K. P. Belov, *Magnetic Transitions* (Fizmatgiz, Moscow, 1959; Consultants Bureau, New York, 1961).

3. S. I. Novikova, *Thermal Expansion of Solids: A Handbook* (Nauka, Moscow, 1974).

4. A. I. Zakharov, *Physics of Precision Alloys with Special Thermal Properties* (Metallurgiya, Moscow, 1986).

5. V. L. Sedov, *Antiferromagnetism of Gamma Iron: The Invar Problem* (Nauka, Moscow, 1987).

6. E. Z. Valiev, Usp. Fiz. Nauk **161** (8), 87 (1991) [Sov. Phys. Usp. **34**, 685 (1991)].

7. A. Yu. Romanov and V. P. Silin, Zh. Éksp. Teor. Fiz. **113**, 213 (1998) [JETP **86**, 120 (1998)].

8. *Precision Alloys: A Handbook*, Ed. by B. V. Molotilov (Metallurgiya, Moscow, 1983).

9. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 1.

10. L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 4th ed. (Nauka, Moscow, 1995), Part 1.

11. V. Yu. Bodryakov, A. A. Povzner, and O. G. Zelyukova, Fiz. Tverd. Tela (St. Petersburg) **40**, 1581 (1998) [Phys. Solid State **40**, 1433 (1998)].

12. V. Yu. Bodryakov, A. A. Povzner, and O. G. Zelyukova, Metally, No. 2, 79 (2000).

13. V. Yu. Bodryakov, V. V. Petrushkin, and A. A. Povzner, Fiz. Met. Metalloved. **89** (4), 5 (2000).

14. V. Yu. Bodryakov and A. A. Povzner, Fiz. Met. Metalloved. **89** (6), 21 (2000).

15. V. Yu. Bodryakov, V. V. Petrushkin, and A. A. Povzner, Fiz. Met. Metalloved. **89** (4), 5 (2000).

16. V. Yu. Bodryakov and A. A. Povzner, *Self-Consistent Thermodynamic Model of Crystal Lattice* (GOU VPO UGTU–UPI, Yekaterinburg, 2002), Chap. 1.

17. V. Yu. Bodryakov and A. A. Povzner, Zh. Tekh. Fiz. **73** (7), 136 (2003) [Tech. Phys. **48**, 931 (2003)].

18. M. Rosen, Phys. Rev. **174**, 504 (1968).

19. V. Yu. Bodryakov, A. A. Povzner, and S. A. Nikitin, Eur. Phys. J. B **4**, 441 (1998).

20. V. Yu. Bodryakov, A. A. Povzner, and O. G. Zelyukova, Fiz. Met. Metalloved. **87** (4), 13 (1999).

21. V. Yu. Bodryakov and A. A. Povzner, Fiz. Met. Metalloved. **89** (5), 15 (2000).

22. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991; CRC, Boca Raton, 1997).

Translated by V. Isaakyan

OPTICS,
QUANTUM ELECTRONICS

Excimer Source of Radiation for Photobiology

A. N. Malinin*, A. V. Polyak*, I. V. Blonskiy**, and N. G. Zubrilin**

* Uzhgorod National University, Podgornaya ul. 46, Uzhgorod, 88000 Ukraine
e-mail: mal@univ.uzhgorod.ua
** Institute of Physics, National Academy of Sciences of Ukraine, pr. Nauki 144, Kiev, 03028 Ukraine
e-mail: zubrilin@iop.kiev.ua
Received June 19, 2003

Abstract—A gas-discharge excimer source of visible radiation employing a mixture of mercury diiodide vapor, mercury dibromide vapor, and helium is studied. The emission spectrum of the source covers the range from 370 to 510 nm. About 90% of the radiation power is concentrated in the blue–green spectral region. © 2004 MAIK "Nauka/Interperiodica".

INTRODUCTION

At present, artificial lighting needed for plant growth is provided by high-pressure sodium-vapor lamps. Only one-third of the radiation power from such sources is concentrated in the spectral range of photosynthetic activity. New-generation radiation sources with selective characteristics in the spectral ranges of 400–510 and 610–720 nm are required for more efficient optical control over photosynthesis and growth of plants and algae.

EXPERIMENTAL

The design of an excimer lamp suggested in this paper is shown in Fig. 1. The diameter and length of quartz tube *2* of the lamp are 34 and 200 mm, respectively. Inner cylindrical tungsten electrode *4* with a diameter of 4 mm is placed along the axis of the tube. The transmission coefficient of second (perforated) electrode *3* is 72%. The power supply is connected to the inner electrode via metal–quartz bushing *6*. Inlet connection *1* made of quartz glass is located on the opposite end face. It is used to evacuate the tube or to fill it with the mixture components.

The working mixture of the excimer lamp is excited in the barrier discharge plasma. The discharge is initiated by a nanosecond pulser. A TGI 2-130/10 thyratron is used as the switch of the pulser. The storage capacitor of the pulser is made up of a set of KVI-3 low-inductance capacitors and is charged through the primary winding of a step-up transformer with a transformation ratio of 1 : 3. In experiments, voltage and current pulses (of duration ≈150 ns) applied to the lamp electrodes were 22–30 kV and 265 A, respectively. The pulse repetition rate was 2–5 kHz, and the capacitance of the storage capacitor was 1.36 nF. The electric parameters of the lamp (voltage and current) were measured with a calibrated voltage divider and Rogowski loop, respectively.

The spectral and integral characteristics of the excimer lamp were studied in the following way. The radiation directed along the normal to the lamp surface passed through a diaphragm of area 1 cm^2, was projected by a lens onto the entrance slit of a ZMR-3 prism monochromator, and detected by an FÉU-79 PMT equipped with a KSP-4 x–y recorder (the detection scheme is similar to that used in [1]). At a wavelength of 434 nm, the inverse linear dispersion of the monochromator is 44 Å/mm. During spectral measurements, the slit width was kept at 0.1 mm. In taking the integral characteristics (the dependence of the spectral line intensity on the mixture composition and the partial pressures of the components), the slit width was 1.5 mm. The detecting system was calibrated with an SI 8-200 standard tungsten lamp at filament temperature $T = 2173$ K.

The pulsed and mean powers of the radiation were measured by means of an FÉK-22 SPU photoelectric cell and a Kvarts-01 device, respectively. They were used in the detecting system instead of the ZMR-3 monochromator.

The radiation power from the entire surface of the excimer source was determined from an expression for irradiance due to an emitting line of equal brightness provided that the length of the line is much less than the
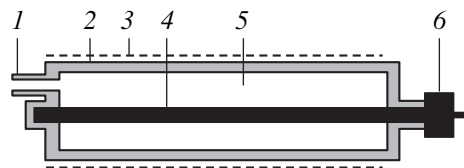

Fig. 1. Design of the excimer lamp: (*1*) inlet connection for evacuation and filling with gas, (*2*) quartz tube, (*3*) perforated electrode, (*4*) electrode, (*5*) discharge space, and (*6*) high-voltage bushing.

photodetector–source distance [2]:

$$P_1 = \Omega_0 P_2 / \Omega_2.$$

Here, P_2 is the power measured by the photodetector, Ω_0 is the equivalent solid angle, $\Omega_2 = S/l_0^2$ is the solid angle of the photodetector, S_2 is the surface area of the photodetector window, and l_0 is the photodetector–source distance.

When calculating the pulsed and mean powers, we took into account the emission area. For the cylindrical surface, the equivalent solid angle equals π^2 [2].

The working mixtures were prepared directly in the radiation source. First, we place equal amounts (60 mg) of mercury diiodide (HgI_2) and mercury dibromide ($HgBr_2$) into the tube and degassed it by heating at 50°C and subsequently evacuating for 2 h. Then, we filled the tube with an inert gas (helium). The partial pressures of the HgI_2 and $HgBr_2$ vapors were set by heating the working mixture via dissipation of the pulsed–periodic discharge energy. The partial pressures were measured from the temperature of the coolest point in the source, which was found by linearly interpolating reference data from [3].

OPTICAL AND ENERGY CHARACTERISTICS

Figure 2 shows the panoramic spectrum of the excimer source containing the vapors of mercury dihalides (mercury diiodide and mercury dibromide) and helium (HgI_2 : $HgBr_2$: He = 14.6 Pa : 49.2 Pa : 162 kPa). The pump pulse repetition rate is 4 kHz; the voltage amplitude, 25.5 kV; and the current, 265 A. The spectra exhibit overlapping emission bands peaked at $\lambda = 444$ and 502 nm (by calibration, the intensity radiation ratio between HgI* and HgBr* molecules equals 2.5), a sharp increase in the intensity in the long-wave part of the spectrum, and a smooth decrease in the short-wave range. With regard to the tails of the spectral bands, the emission covers the range from 370 to 510 nm. As the repetition rate of pump pulses varies from 2 to 5 kHz, the intensities of the spectral bands and the intensity ratio at their edges change, whereas the shape and range of the spectral bands, as well the positions of the maxima, remain unchanged.

When the partial pressure of helium rises from 140 to 200 kPa, we observe a nonmonotonic variation of the mean radiation power P_3: it first grows in the interval 140–180 kPa, reaches a maximum at 182 kPa, and then drops (Fig. 3). Figures 4 and 5 show that P_3 increases linearly with the voltage and the pulse repetition rate. For a voltage of 30 kV, $P_3 = 11$ W for a pulse repetition rate of 2 kHz and ≈ 22.5 W at a pulse rate of 5 kHz.

When the partial pressures of the mercury dihalides increase, P_3 reaches a maximum at a HgI_2 pressure of 0.25 kPa and a $HgBr_2$ pressure of 0.5 kPa. The dependence of P_3 on the partial pressures of HgI_2 and $HgBr_2$
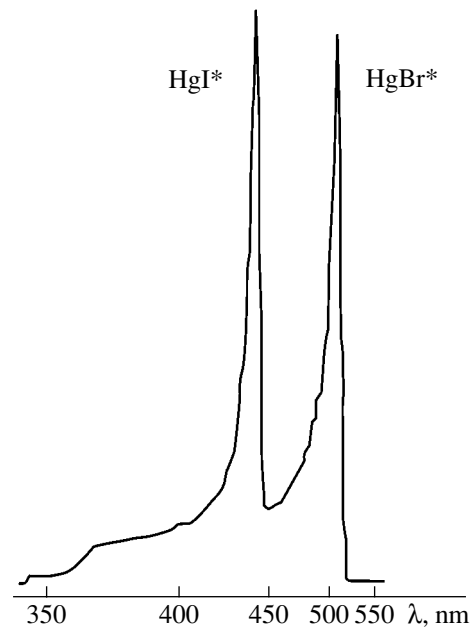
Fig. 2. Radiation spectrum of the excimer source with the HgI_2 : $HgBr_2$: He = 14.6 Pa : 49.2 Pa : 162 kPa working mixture at a pump pulse repetition rate of 4 kHz, a voltage of 25.5 kV, and a current of 265 A.
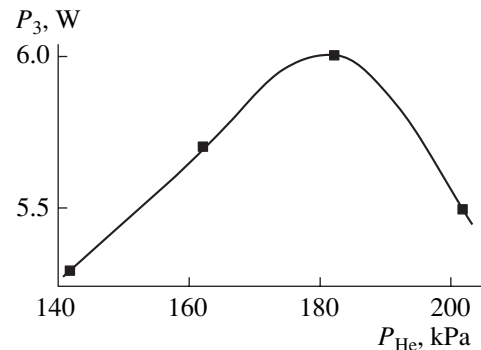
Fig. 3. Radiation mean power vs. the partial pressure of helium for a pulse repetition rate of 2 kHz and a voltage amplitude of 22.5 kV.
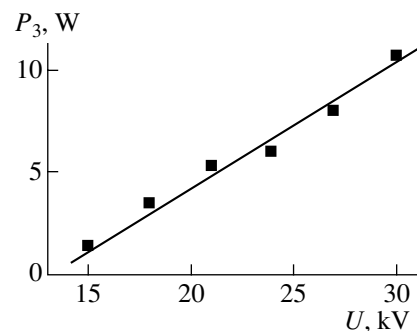
Fig. 4. U dependence of the radiation mean power. The total pressure of the mixture is 162 kPa, and the pulse repetition rate is 2 kHz.
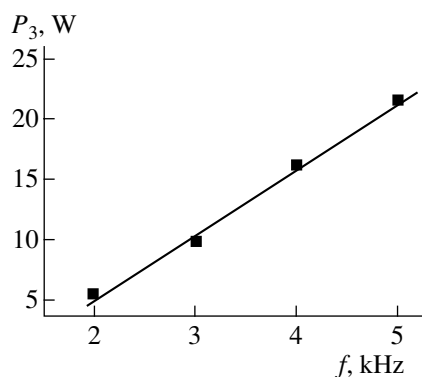
Fig. 5. Radiation mean power vs. the pulse repetition rate. The total pressure of the mixture is 162 kPa, and the voltage amplitude is 22.5 kV.
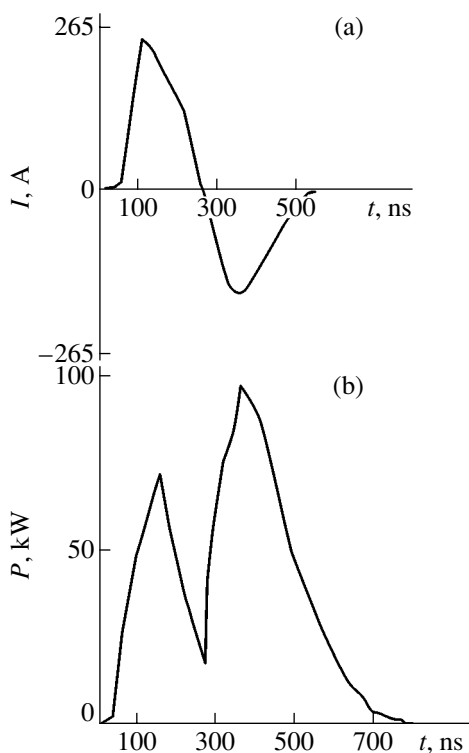

Fig. 6. Oscillograms of (a) the discharge current and (b) radiation power. The total pressure of the mixture is 162 kPa, the pulse repetition rate is 2 kHz, and the voltage amplitude across the lamp is 22.5 kV.

was taken by heating the working mixture with an external electric heater.

Figure 6 demonstrates typical waveforms of the discharge current and radiation power for a total pressure of the mixture of 162 kPa, a voltage of 22.5 kV, and a pulse repetition rate of 2 kHz. The accuracy and reproducibility of the waveforms are 10 and 90%, respectively. The amplitude and duration of the bipolar current pulses applied were 265 A and 150 ns, respectively.
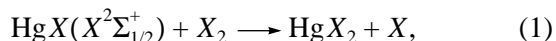
It is seen that $P_3(t)$ is a double-humped function and the peak values of the current and power coincide in time. The amplitude of the second radiation pulse is greater than that of the first one, although the amplitude of the second current pulse is lower than the amplitude of the first one. In addition, the overall duration and the trailing edge time of the second radiation pulse are longer than the corresponding parameters of the first pulse.
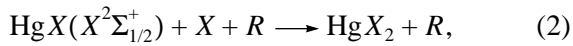
The emission bands shown in Fig. 2 correspond to the electronic–vibrational transitions $B^2\Sigma^+_{1/2} \longrightarrow X^2\Sigma^+_{1/2}$ in HgI* and HgBr* molecules [4]. A change in the pulse repetition rate changes the temperature of the lamp wall. Therefore, the associated changes in the partial pressures (in the concentrations of the mercury diiodide and dibromide vapors) will be different [3]; hence, the concentrations of HgI* and HgBr* molecules in the excited state $B^2\Sigma^+_{1/2}$ will also differ [5]. This is embodied in the different intensities of the emission bands.

The optimum pressure of the buffer gas (helium) (Fig. 4) is related to the discharge energy spent on heating the mixture of the two mercury dihalides and helium. This energy depends on the parameter E/p (the ratio of the field strength in the plasma to the pressure of the mixture). For binary mixtures (the vapor of one of the mercury dihalides plus helium), it was demonstrated [6] that there exists an optimum range of wall temperatures where the radiation intensity is maximum. It is known [7] that, when the wall temperature exceeds a certain value, the quenching of the state $B^2\Sigma^+_{1/2}$ of these molecules starts playing a significant role.

The variation of P_3 with partial pressures of the mercury dihalides is explained both by the increase in the molecular concentration in the state $B^2\Sigma^+_{1/2}$ with partial pressures and by the quenching of these states. The optimum partial pressures of the mercury dihalides are established when these two processes are in dynamic equilibrium.

The time dependence of the current is a result of charge exchange in the insulator–plasma circuit. The presence of two humps with a higher second hump was reported earlier [1, 8] when we studied the time variation of the radiation intensity from HgBr* and HgI* excimer molecules in mercury diiodide–helium and mercury dibromide–helium binary mixtures [1, 8]. The broadening of the second pulse and the extension of its trailing edge in the three-component mixture in comparison with the two-component mixtures may be related to a decrease in the rate of association of mercury monohalide diatomic molecules in the reactions

$$\mathrm{Hg}X(X^2\Sigma^+_{1/2}) + X_2 \longrightarrow \mathrm{Hg}X_2 + X, \qquad (1)$$

$$\mathrm{Hg}X(X^2\Sigma^+_{1/2}) + X + R \longrightarrow \mathrm{Hg}X_2 + R, \qquad (2)$$

where X = Be or I and R is a buffer gas. Accordingly, the number of mercury monohalide molecules in the ground state grows. Eventually, the waveform of the second pulse changes, since the excitation of the state $X^2\Sigma^+_{1/2}$ increases the electron population of the state $B^2\Sigma^+_{1/2}$ of mercury monohalides [6].

CONCLUSIONS

Thus, we studied an excimer source of visible radiation that contains mercury diiodide vapor, mercury dibromide vapor, and helium as a working mixture and is excited by a pulsed–periodic barrier discharge. It was demonstrated that the spectrum of the source consists of overlapping spectral bands in the wavelength range 370–510 nm, which correspond to emission due to mercury monoiodide and mercury monobromide molecules. The mean power of radiation from the active area of the source (230 cm^2) is about 25 W, and the peak power equals 93 kW. The discharge-to-radiation power conversion efficiency is 30%. About 90% of the radiation power is concentrated in the blue–green spectral range. This value is significantly higher than the corresponding parameter of high-pressure sodium-vapor lamps for comparable energy deposits in the working medium [9].

The results obtained suggest that optimization of the excitation scheme and the design of the lamp may improve the energy parameters of the excimer lamps. This source may provide a more effective control of photosynthesis and growth of plants and algae.

REFERENCES

1. A. N. Malinin, A. V. Polyak, N. G. Zubrilin, *et al.*, Kvantovaya Élektron. (Moscow) **32**, 155 (2002).
2. R. A. Sapozhnikov, *Theoretical Photometry* (Énergiya, Moscow, 1977).
3. A. I. Efimov, L. P. Belorukova, I. V. Vasil'kova, and V. P. Chechev, *Properties of Inorganic Compounds* (Khimiya, Leningrad, 1983).
4. R. W. Pears and A. G. Gaydon, *The Identification of Molecular Spectra* (Chopman Hall, London, 1963).
5. A. N. Malinin, Laser Phys. **8**, 395 (1998).
6. A. N. Malinin, Laser Phys. **7**, 1032 (1997).
7. A. N. Malinin, Laser Phys. **7**, 1177 (1997).
8. A. N. Malinin, A. K. Shuaibov, and V. S. Shevera, Zh. Prikl. Spektrosk. **32**, 735 (1980).
9. F. G. Baksht and V. F. Lapshin, Zh. Tekh. Fiz. **67** (9), 22 (1997) [Tech. Phys. **42**, 1004 (1997)].

Translated by A. Chikishev

OPTICS,
QUANTUM ELECTRONICS

Hysteresis of the Spectral Components of a Femtosecond Pulse Propagating in a Nonlinear Medium

D. K. Skripov and V. A. Trofimov

Moscow State University, Vorob'evy Gory, Moscow, 119992 Russia
e-mail: vatro@cs.msu.su
Received July 8, 2003

Abstract—Numerical simulation based on the one-dimensional nonlinear Maxwell equations is used to study the frequency that provides the maximum spectral intensity in a selected time interval versus the amplitude of an incident femtosecond pulse. The hysteretic variation of this frequency with amplitude is shown to be a possibility. Analysis is performed in the approximation of an optically thick layer for a medium with a saturating restoring force. © 2004 MAIK "Nauka/Interperiodica".

INTRODUCTION

Femtosecond optics is finding ever increasing application in various fields of science and engineering. The short length of these pulses allows researchers to study various ultrafast processes. At the same time, this feature causes hitherto unknown specific interactions between an optical pulse and a medium. An example is the variation of the electromagnetic wave spectrum with the absolute phase of a femtosecond pulse at the input to a nonlinear medium. In particular, this effect was found in the direct solution of the nonlinear Maxwell equations [1]. The effect of the absolute phase on the dynamics of interaction between a femtosecond pulse and the medium was experimentally demonstrated in [2].

Another intriguing effect important for applications is the hysteretic dependence of the frequency providing the maximum spectral intensity on the signal amplitude. For an optically thin medium, this effect is described in [3, 4]. In this paper, we study the possibility for such behavior in thick media.

BASIC EQUATIONS

Consider the propagation of a femtosecond pulse along the z axis in terms of the nonlinear one-dimensional Maxwell equations for the electric, $E(z, t)$, and magnetic, $H(z, t)$, fields with a saturating restoring force in the case when the polarization $P(z, t)$ of the medium is analyzed:

$$\frac{\partial H}{\partial z} = -\frac{\partial D}{\partial t}, \quad \frac{\partial E}{\partial z} = -\frac{\partial H}{\partial t},$$

$$D = E + 4\pi P, \quad -L_i < z \le L_z, \quad 0 < t \le L_t, \quad (1)$$

$$\frac{\partial^2 P}{\partial t^2} + \delta\frac{\partial P}{\partial t} + \frac{P}{1 + \left(\frac{P}{P_0}\right)^n} = \alpha E(z, t), \quad z \ge 0.$$

We assume that the nonlinear medium is to the left of the section $z = 0$. Therefore, $P = 0$ for $z < 0$.

In Eqs. (1), z is the spatial coordinate; L_i is the length of the region in contact with the nonlinear medium; L_z is the longitudinal dimension of the nonlinear medium; and t and L_t are the dimensionless time and its maximum value, respectively. The coefficient δ characterizes the attenuation of the polarization of the medium, and α is a factor proportional to the dipole moment of an atom or molecule. The variables are normalized so that the linear oscillator frequency equals unity. The parameter P_0 makes computer simulation more convenient. The parameter n is set equal to $n = 4$, which means field–dipole interaction [5]. It should be emphasized that the potential in the form of a saturating function is widely met in the literature. For example, the dipole moment in the form $re^{-\frac{r}{r_0}}$ was used in [6] to describe interaction between laser radiation and a molecule. It is also important to note that we consider such intensities of light that do not ionize an atom.

The initial and boundary conditions for Eqs. (1) are specified in the form

$$E|_{z = -L_i} = H|_{z = -L_i} = P|_{t = 0} = \frac{\partial P}{\partial t}\bigg|_{t = 0} = 0,$$

$$E|_{t = 0} = E_0(z), \quad H|_{t = 0} = H_0(z), \quad (2)$$

where the initial electric field distribution in the optical wave is defined outside the nonlinear medium (i.e., in the linear medium) within the segment $[-L_i, 0]$ in the form of a triangular pulse

$$E(0, z) = E_0(z)$$

$$= \begin{cases} E_0(z)\cos(\omega(z + 0.5L_i)), \\ E_0(z) = E_0(1 - |1 + 2z/L_i|), & -L_i \le z < 0, \\ 0, & 0 \le z < L_z, \end{cases} \quad (3)$$

$$H_0(z) = E_0(z).$$

The initial distribution of the magnetic field H is chosen to be the same as that of the electric field. In this situation, the electromagnetic field propagates in a homogeneous medium along the positive z direction. In formula (3), ω is the filling (carrier) frequency of the incident pulse and E_0 is the pulse amplitude. We consider a pulse that totally occupies the linear medium $z < 0$; then, L_i is the characteristic length of this medium. The choice of the incident pulse shape stems from the goal of our work: the triangular waveform illuminates the behavior associated with optical bistability. Such waveforms are commonly used in related analysis.

To find the spectral response of the medium to the instantaneous amplitude of the incident triangular pulse, we will use the following algorithm. Let us divide the pulse duration $[0, L_t]$ into M equal intervals of length T ($t_k = kT$, $k = 0\text{–}M$, $L_t = MT$) and apply Fourier transformation to the electric induction on each of the intervals $[t_k, t_{k+1}]$:

$$D_k(\omega, z) = \frac{1}{T} \int_{t_k}^{t_k + T} D(t + z, z)e^{-i\omega(t - t_k)}dt, \quad (4)$$

$$k = 0, 1, \ldots, M - 1.$$

The function $D(t, z)$ is integrated over the interval $t \in [z + t_k, z + t_k + T]$, since part of the leading edge of the pulse propagates with a velocity equal to unity (linear precursor) because of the fact that the medium responds with a time delay. Then, the pulse arrives at a section z at the time instant $t = z$. Therefore, a distortion of the pulse at the entrance to the medium at the time t arrives at the section z at the time instant $t + z$.

Accordingly, the inverse Fourier transform has the form

$$D(t + z, z) = \int_{-\infty}^{\infty} D_k(\omega, z)e^{i\omega(t - t_k)}d\omega, \quad t \in [t_k, t_{k+1}]. \quad (5)$$

In our numerical experiments, we use the fast discrete Fourier transform instead of formulas (4) and (5)

and define the spectral mode intensity as the magnitude of the corresponding harmonic squared:

$$I_k(\omega, z) = |D_k(\omega, z)|^2. \quad (6)$$

In formula (4), a part of the spatial interval that corresponds to $k = M - 1$ (namely, $[L_t, L_t + z]$) is outside the time interval of interest. Therefore, the function $D(t + z, z)$ is integrated over the interval $L_t - T \le t + z \le L_t$; i.e., the interval of the Fourier transform is shifted by z towards zero until it falls into the time domain of interest. This affects the results of simulation only slightly for several reasons. First, after the pulse has passed, polarization oscillations small in amplitude and almost constant in frequency and amplitude are observed at the edge of the time domain. Second, the time shift by z is small compared with the length of each of the kth intervals. Third, in our numerical experiments, the maximum frequency is kept constant in the last several intervals. Also note that the scenario described in this paper is observed near the center of the incident pulse.

It should be emphasized that such analysis of the instantaneous spectral content of a waveform is widely used in the theory of signal processing [7]. To this end, a "window" of constant length travels along the waveform and the spectrum within the window is analyzed. In our case, the windows follow each other in time without overlap (except for the next to last and last intervals, in view of the above comment).

The appropriate difference schemes and numerical methods are reported in [8].

RESULTS OF COMPUTER SIMULATION

It is clear that the nonlinearity of the medium generates new spectral lines of various intensities. Therefore, in each time interval and in a particular section, several spectral lines with center frequencies $\omega_{k,m}$ may be
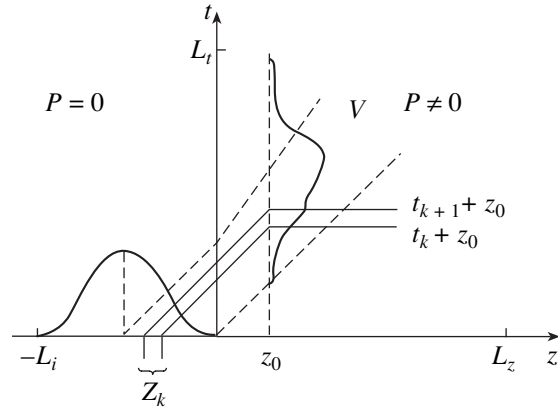

Fig. 1. Correspondence between the electric field induction in a medium and electric field intensity at the input to the medium in the calculation of the frequency providing the maximum spectral intensity as a function of the incident radiation amplitude.
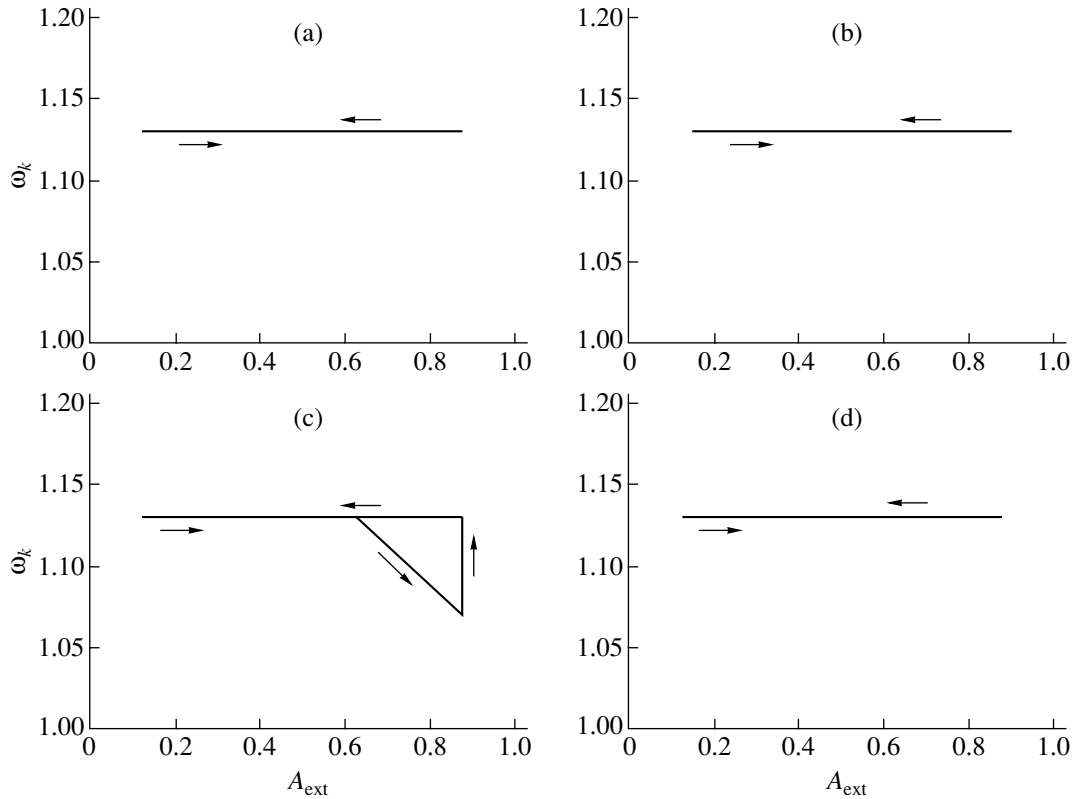
Fig. 2. Hysteretic dependence of the frequency of the highest intensity harmonic on the averaged amplitude of the triangular pulse propagating through the (a, c) nonlinear and (b, d) linear media in the sections z = (a, b) 0, 1, 2 and (c, d) 4, 6 with $\alpha = 0.01$, $\omega = 1.1$, $\delta = 10^{-3}$, $P_0 = 0.15$, $E_0 = 1$, $L_i = L_z = L_t = 800$, and $T = 100$. The power reflection coefficient is 4.7% for nonlinear propagation and 5.1% for linear propagation.

present (the subscript k stands for the time interval considered, and the subscript m indicates a local spectral maximum in this time interval). In going from one time interval to another and as the pulse propagates through the medium, these maxima may move, disappear, and change amplitude. As a result, the frequency corresponding to the maximum spectral intensity in a particular time interval and in a particular section may change. Therefore, in our numerical experiments, we choose the frequency that provides the maximum intensity in a section z for the majority of the time intervals; in other words, the global maximum of the spectrum is chosen. In those intervals where the maximal spectral intensity is reached at a much different frequency, this frequency is taken to be the frequency of the global maximum.

Below, the results of our calculations are presented graphically as dependences of the frequency ω_k,

$$\omega_k = \max_{\omega} I_k(\omega, z), \quad k = 0-(M-1), \qquad (7)$$

on the external perturbation amplitude averaged over a given interval:

$$A_{ext} = (E_0|_{z = -t_{k+1}} + E_0|_{z = -t_k})/2. \qquad (8)$$

Figure 1 schematically shows the correspondence between the electric induction and initial electric field distribution. Since the response is nonstationary and nonlinear, its velocity in the coordinates adopted is characterized by the slope of the line along which the wave propagates.

For convenience, the direction of time variation is shown in Figs. 2–4 by arrows. An increase in A_{ext} refers to the leading edge of the pulse; a decrease in A_{ext}, to its trailing edge. Since the nonstationarity of the response may significantly affect the hysteretic behavior, Figs. 2–4 also show the results for a linear nonstationary medium.

Figures 2–4 refer to different excitations of the medium. In particular, in Fig. 2, the medium is excited weakly and its contribution to the electric induction is insignificant. Here, the nonlinear medium reflects about 5% of the incident power and the rest of the power penetrates into it. In this case, the variation of the frequency that provides the maximal spectral intensity with incident optical pulse amplitude within the thickness of the medium ($z = 4$ and 6 in Fig. 2c) becomes uncertain (hysteretic). For the linear nonstationary response (Fig. 2d), such behavior is not observed. It is also
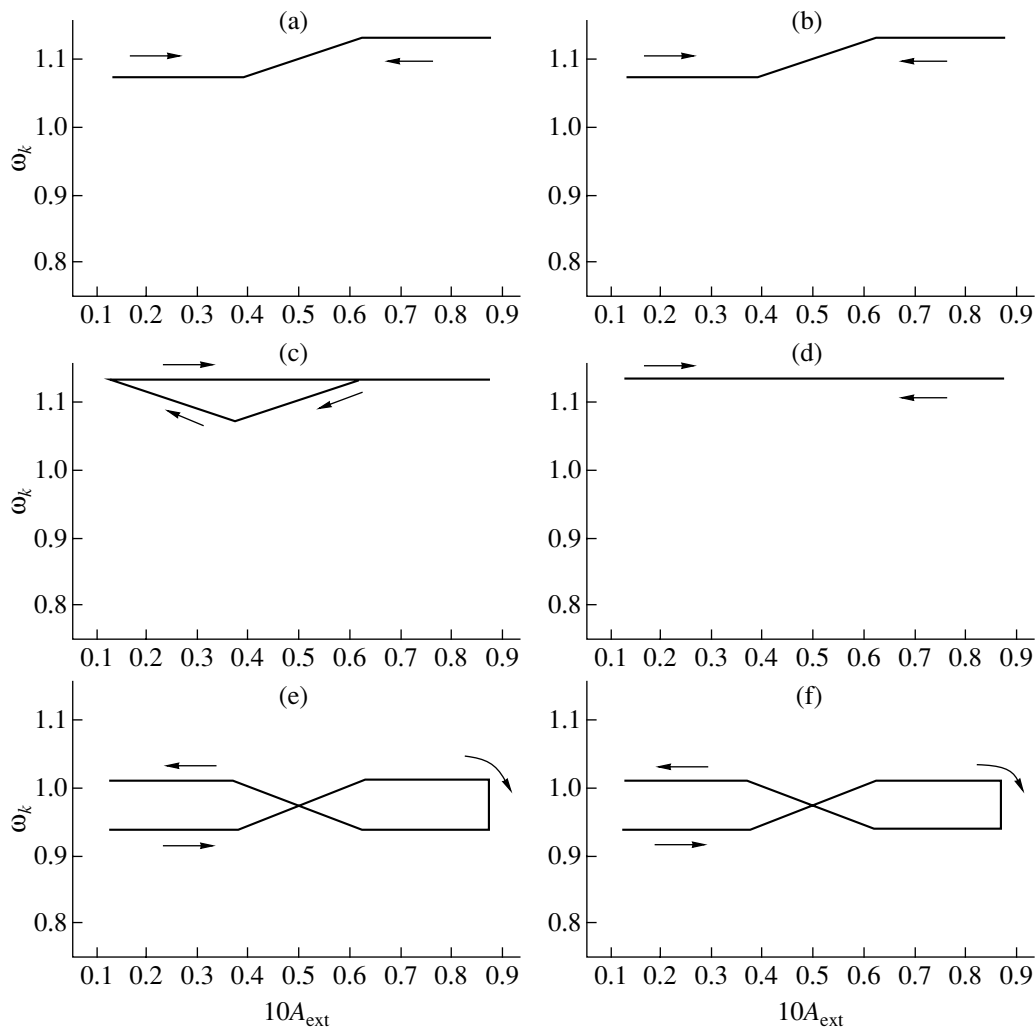
Fig. 3. Hysteretic dependence of the frequency of the highest intensity harmonic on the averaged amplitude of the triangular pulse propagating through the (a, c, e) nonlinear and (b, d, f) linear media in the sections $z = $ (a, b) 0, (c, d) 1, and (e, f) 2 with $\alpha = 0.2$, $\omega = 1.1$, $\delta = 10^{-3}$, $P_0 = 0.1$, $E_0 = 0.1$, $L_i = L_z = L_t = 800$, and $T = 100$. The power reflection coefficient is 99.7%.

important that the above variation remains unique up to $z = 4$.

When the excitation of the medium is strong (Figs. 3 and 4), most of the energy reflects from the nonlinear medium and the ω_k versus input amplitude dependence becomes more complex. For applications, it is important that, all other parameters remaining constant, the duration of the incident pulse is of considerable significance, as illustrated in Figs. 3 and 4. They clearly show that doubling of the pulse duration changes qualitatively the dependence of the frequency providing the maximal intensity on the pulse amplitude. In the case of the longer pulse, the effect of nonstationarity on the frequency ω_k dominates over that of nonlinearity. Indeed, while Fig. 3c demonstrates a clear hysteretic behavior, a similar dependence that differs from the linear nonstationary response is not observed in Fig. 4.

Several features are worth noting. First, the hysteretic curves considered above depend significantly on the incident pulse waveform. In particular, closed hysteresis loops were not observed for a Gaussian pulse, because this pulse generates intense high-frequency components at its leading edge and low-frequency components prevailing at its trailing edge. A decrease in the Gaussian pulse length enhances the generation of high-frequency components. Second, with an increase in the frequency mismatch between the incident pulse and the linear resonance of the system, the hysteresis loops disappear (all other parameters remaining unchanged). The response of the medium either becomes essentially nonlinear (i.e., bears no resemblance to that of a linear system) or coincides with the linear response that is found when the evolution of the maximum intensity harmonic is studied by moving the window in the time domain of simulation.
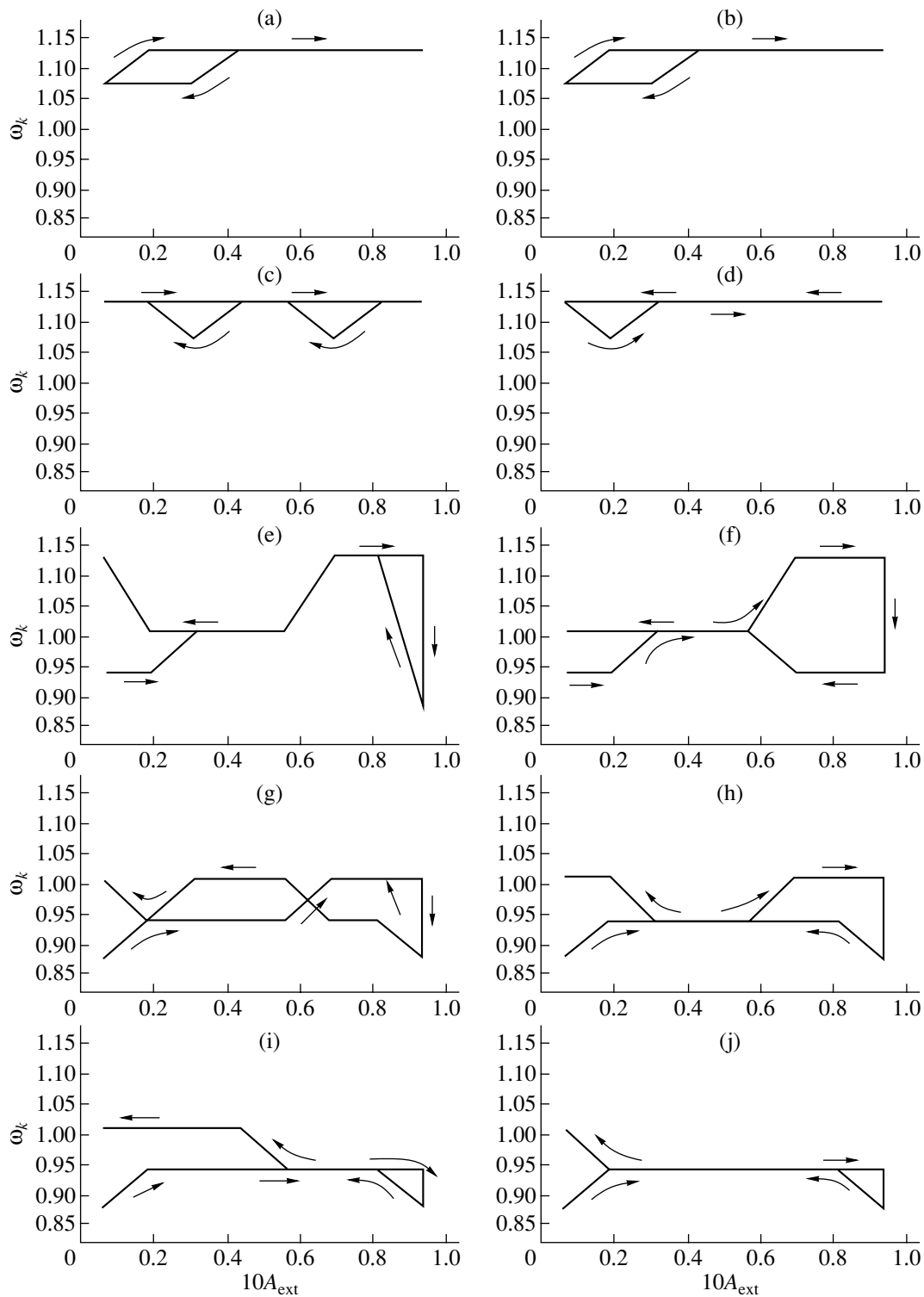
Fig. 4. Hysteretic dependence of the frequency of the highest intensity harmonic on the averaged amplitude of the triangular pulse propagating through the (a, c, e, g, i) nonlinear and (b, d, f, h, j) linear media in the sections $z = $ (a, b) 0, (c, d) 1, (e, f) 2, (g, h) 4, and (i, j) 6 with $\alpha = 0.2$, $\omega = 1.1$, $\delta = 10^{-3}$, $P_0 = 0.1$, $E_0 = 0.1$, $L_i = L_z = L_t = 1600$, and $T = 100$. The power reflection coefficient is 99.7%.

CONCLUSIONS

Thus, when a femtosecond pulse propagates through a nonlinear medium, the variation of the fre-

quency that provides the maximal intensity with the amplitude of the incident pulse may be hysteretic within the thickness of the medium. However, unlike the case of an optically thin layer, this hysteresis is

much more difficult to observe when the medium is thick (i.e., in our case). The appearance of the hysteresis depends significantly on the excitation level, the nonstationarity of the process, and the mismatch between the carrier frequency and the linear resonance frequency. The behavior described above could be observed in physical experiments [2] where the chemical reaction yield or the number of photoelectrons of various energy at the leading and trailing edges of the pulse is different. Nonlinear properties of a medium also allow us to form its response at frequencies that provide different spectral intensities. In other words, we can produce a train of spikes of different repetition rate (at the leading and trailing edges of the applied pulse) directly in a chemically active medium and control the chemical reaction by the technique proposed in [6].

REFERENCES

1. D. K. Skripov and V. A. Trofimov, Pis'ma Zh. Tekh. Fiz. **27** (14), 6 (2001) [Tech. Phys. Lett. **27**, 575 (2001)].
2. G. G. Paulus *et al.*, in *Proceedings of the International Quantum Electronics Conference, Moscow, 2002*, p. 231.
3. V. A. Trofimov and D. K. Skripov, in *Abstracts of the Second Conference "Superstrong Fields in Plasmas," Varenna, Italy, 2001*, p. 8.
4. D. K. Skripov and V. A. Trofimov, Zh. Tekh. Fiz. **73** (3), 69 (2003) [Tech. Phys. **48**, 344 (2003)].
5. S. M. Gladkov and N. I. Koroteev, Usp. Fiz. Nauk **160** (7), 105 (1990) [Sov. Phys. Usp. **33**, 554 (1990)].
6. M. V. Korolkov, J. Manz, and G. K. Paramonov, Chem. Phys. **217**, 341 (1997).
7. K. Chui, *Introduction to Wavelets* (Mir, Moscow, 2001).
8. D. K. Skripov and V. A. Trofimov, Volokonno-Opt. Tekhnol., Mater., Ustr., No. 3, 82 (2000).

Translated by A. Khzmalyan

Resonant Destabilization of Phonon Microwave Stimulated Emission in an Acoustic Quantum Oscillator (Phaser) with Periodically Modulated Pumping

D. N. Makovetskii

Usikov Institute of Radio Physics and Electronics, National Academy of Sciences of Ukraine,
61085 Kharkov, Ukraine
e-mail: makov@ire.kharkov.ua
Received October 31, 2002; in final form, July 16, 2003

Abstract—Two qualitatively different types of resonant destabilization of phonon stimulated emission (SE) are discovered in experiments where a 9-GHz multimode ruby laser is periodically modulated (the electromagnetic pump frequency is 23 GHz). In the case of deep pump modulation at low modulation frequencies (ω_m = 70–200 Hz, where ω_m is the modulation frequency), a fast random alternation of microwave phonon SE modes is observed. This destabilization range corresponds to relaxation resonance in optical lasers. Outside the relaxation resonance range (at $\omega_m \approx 10$ Hz), the other type of resonant destabilization of stationary phonon SE is observed. This destabilization shows up as very slow regular self-detunings of the microwave SE spectra. The period of such self-organized motions depends significantly on ω_m and changes by several orders of magnitude when ω_m varies within several percent. The second type of SE resonant destabilization is explained in terms of antiphase energy exchange between modes in a modulated phaser. © *2004 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

The feasibility of phonon SE in activated crystals was considered as early as in the 1960s [1, 2]. Yet, speculation about mechanisms behind stimulated emission of phonons persists even today (see, e.g., [3]). In experiments, phonon SE was first observed and studied [4–8] in dielectric crystals doped by paramagnetic iron-like ions. The SE effect showed up as the quantum paramagnetic amplification of a coherent microwave phonon flux (hypersound) when spin levels that may take part in spin–phonon interaction were inversely populated. This effect may be viewed as an acoustic analogue [9] of maser amplification of electromagnetic waves (if a number of features of nonlinear processes in the signal and pump channels are taken into account [10–13]).

At the same time, the mechanism of phasing (generation of microwave acoustic radiation), which was discovered experimentally in [14, 15], long remained unclear. The reason was an attempt to draw an analogy [16] between acoustic quantum oscillators (phasers) and electromagnetic quantum microwave oscillators (masers), which does exist between the corresponding amplifiers.

In experimental studies [11, 17, 18] of microwave acoustic SE in $Al_2O_3 : Ni^{2+}$ and $Al_2O_3 : Cr^{2+}$ crystals, it was shown that phasing is physically much closer to lasing than to masing. In fact, the hypersonic wavelength in a Fabry–Perot acoustic resonator (FPAR) is roughly 1–3 μm (i.e., falls into the near-IR range). The quality factor Q_C of an FPAR, as well as the quality factor of electromagnetic cavities in many lasers, is high: $Q_C = 10^5$–10^6 [11, 17, 18] (certainly, this value is reached at liquid helium temperatures, when the nonresonant decay of hypersound is low). Therefore, experimental SE spectra of phasers operating in the autonomous regime [11, 15, 17] sometimes are similar to those observed for class-B multimode solid-state lasers (with $\tau_1 \gg \tau_C \gg \tau_2$, where τ_1 and τ_2 are the relaxation times for longitudinal and transverse relaxations of active centers and τ_C is the lifetime of field excitations in the cavity).

However, phasers differ radically from lasers in regard to the intrinsic quantum noise (spontaneous radiation) intensity J_{spont}. Since the velocity of hypersound v_h is five orders of magnitude lower than the velocity of light, the SE frequency Ω in a phaser with a hypersonic wavelength of 1–3 μm lies in the range $\Omega = 3$–10 GHz [12, 18, 19], i.e., is five orders of magnitude lower than in a laser. Accordingly, the spontaneous radiation intensity in a phaser is ≈15 orders of magnitude lower than in a laser (because J_{spont} grows as Ω^3). In essence, a phaser may be considered as a deterministic dynamic system throughout the SE intensity range available. This is of crucial importance for studying motion in systems with a complex stratified phase space. It is known that multiplicative noise (including spontaneous radiation in a nonlinear active medium) affects the behavior of dynamic systems in a very intriguing manner [19], causing coarsening of the phase space topology [20], etc.

Earlier [21], we discovered a severe dynamic contraction of the SE spectra in a nonlinear FPAR, which was attributed to the resonant destabilization of mode energy exchange. This noise-unrelated effect also changes qualitatively the phase space topology. Further investigation of this microwave deterministic system as a part of a nonautonomous phaser allowed us to reveal a still more unexpected property: slow large-scale laminar self-detunings of the SE spectra akin to autowave motion [18, 22]. Below, we report these experimental findings and treat them theoretically.

1. EXPERIMENTAL

1.1. Fabry–Perot Microwave Acoustic Resonator, Active Centers, and Hypersonic Converter

Experiments were carried out by using a ruby phaser [11, 21, 23] with the pump power P periodically modulated at low and ultralow frequencies: $\omega_m/2\pi = 1$ Hz–3 kHz (hereafter, the factor $1/2\pi$ will be omitted). A solid-state FPAR, which was made of $Al_2O_3 : Cr^{3+}$ single-crystalline pink ruby, had the form of a cylinder with the diameter $d_C = 2.6$ mm and length $L_C = 17.6$ mm. The end faces of the cylinder are parallel to each other and optically smooth: they serve as acoustic mirrors for hypersonic waves. The triad axis $\mathbb{O}_3$ of the ruby coincides with the geometrical axis $\mathbb{O}_C$ of the FPAR. The concentration of Cr^{3+} ions is $C_a = 1.3 \times 10^{19}$ cm^{-3} ($\approx 0.03\%$). All measurements were made in the interval $T = 1.8$–4.2 K.

For a hypersonic frequency near $\Omega = 9.1$ GHz and the L_C value mentioned above, the separation between longitudinal acoustic modes of the FPAR is $\Delta\Omega_N^0 \equiv \Omega_N^{(0)} - \Omega_{N-1}^{(0)} = 310$ kHz. Here, $\Omega_N^{(0)}$ is the frequency of an Nth mode of the FPAR in the "cold" regime, i.e., at $P = 0$. The frequencies of the hypersound emitted, i.e., the frequencies of phonon SE modes ($\Omega_N \approx \Omega_N^{(P)}$) in the "hot" regime ($P > P_{ph}$, where P_{ph} is the pump power at which phasing starts), lie near 9.12 GHz according to the frequency $\Omega_S = \Omega_{32} \equiv \hbar^{-1}[E_3(H) - E_2(H)]$ of the inverted spin transition between active centers $E_3 \rightleftarrows E_2$ in a static magnetic field $H \approx H_0$. Thus, the frequency Ω_S corresponds to the vertex of the acoustic paramagnetic resonance (APR) line coincident (along the magnetic field) with two electron spin resonance (ESR) lines for the pump. The value of H_0 depends on the frequency Ω_P of the pump microwave field, which saturates the spin transitions $E_1 \rightleftarrows E_3$ and $E_2 \rightleftarrows E_4$ with $\Omega_P = \Omega_{31} = \Omega_{42} \gg \Omega_S$, where $\Omega_{31} = \hbar^{-1}[E_3(H_0) - E_1(H_0)]$ and $\Omega_{42} \equiv \hbar^{-1}[E_4(H_0) - E_2(H_0)]$ (Fig. 1). The symbols $|\psi_i\rangle$ denote wave functions that belong to the energy levels E_i of the ground spin quadruplet (spin $S = 3/2$, orbital quantum number $L = 0$) of a Cr^{3+} ion in the crystal field of ruby. Since this field is of trigonal symmetry,
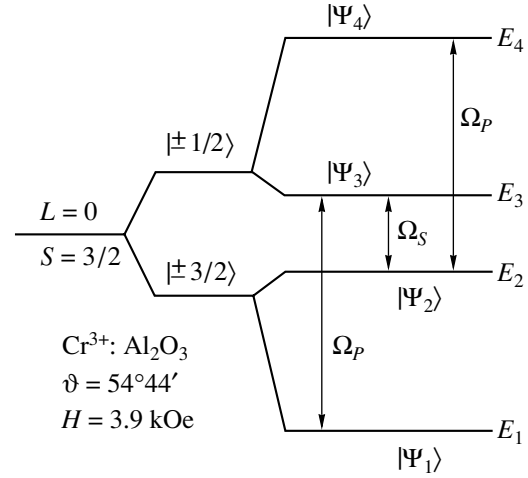
Fig. 1. Energy levels of Cr^{3+} active centers in a ruby phaser.

all E_i and $|\psi_i\rangle$ depend only on $H \equiv |\mathbf{H}|$ and the azimuthal angle ϑ between $\mathbf{H}$ and $\mathbb{O}_3$ [24].

One of the mirrors was covered by a thin (approximately 0.5 µm thick) textured ZnO piezoelectric film with a 0.1-µm-thick Al sublayer (both layers were applied by vacuum evaporation). The texture axis runs normally to the FPAR mirror. The ZnO film is the basic component of a bidirectional hypersonic converter designed for converting a microwave phonon field to an electromagnetic field and vice versa. The phonon SE arising in the FPAR excites electromagnetic oscillations in the ZnO film, and the electromagnetic signal may be detected by standard microwave techniques. On the other hand, exciting the ZnO film from the outside by electromagnetic waves with a frequency Ω_S, we inject hypersonic waves with the same frequency into the FPAR, with $\lambda_h^{(S)} \approx 3.3 \times 10^{-4}\lambda_e^{(S)}$ (where $\lambda_h^{(S)} \approx 1$ µm is the wavelength of longitudinal hypersound in our system and $\lambda_e^{(S)} \approx 3$ cm is the wavelength of an electromagnetic wave of the same frequency as the hypersound).

1.2. Inverse States of Active Centers and Phasing Self-Excitation Conditions

As was noted, inverse spin states of Cr^{3+} active centers are formed by the pump microwave electromagnetic field. The frequency Ω_P of the pump may be tuned within 22–25 GHz; that is, the pump wavelength $\lambda_e^{(P)}$ is near 1.25 cm. The pump power reaches a maximum, $P = P^{(max)} = 12$ mW, at $\Omega_P \approx 23$ GHz. Through a diffraction coupler, the pump field is excited in a cylindrical electromagnetic pump cavity of type H_{011}, which has an eigenfrequency $\Omega_{CP}^{(0)} = 23.0$ GHz, a quality factor $Q_{CP} \approx 8 \times 10^3$, and a geometrical length coincident with the length L_C of the FPAR.

The ruby FPAR is placed in the pump cavity along its axis. If $P = 0$ and the magnitude and direction of $\mathbf{H}$ are beyond the APR range, the absorption of the hypersound injected into the FPAR depends on (i) the nonresonant volume attenuation η_{vol} (including losses on the lateral surfaces of the FPAR) and (ii) losses on the FPAR mirrors η_{mirr}.

If $P = 0$ and the magnitude and direction of $\mathbf{H}$ fall into the APR range (i.e., $\mathbf{H} \approx \mathbf{H}_0$), a third absorption mechanism comes into play: the resonant paramagnetic absorption of the hypersound, which depends considerably on the frequency of the signal injected and on the offset of the magnetic field from the APR line vertex [25].

Finally, $\mathbf{H} \approx \mathbf{H}_0$ and the pump power is applied. Then, the resonant paramagnetic absorption of the hypersound decreases. If, as P rises, one succeeds in passing into the range where the paramagnetic absorption becomes negative (i.e., the inversion ratio $K(P, \mathbf{H})$ becomes positive), nonparamagnetic losses of the hypersound in the FPAR are compensated partially or completely. The complete compensation of the losses (i.e., the onset of phasing) takes place first at that mode (let its frequency be Ω_1) closest to the center of the APR inverted line, for which the condition

$$\frac{1}{Q_{vol}^{(1)}} + \frac{1}{Q_{mirr}^{(1)}} + \frac{1}{Q_{magn}^{(1)}} < 0 \qquad (1)$$

is met prior to other modes. In (1), $Q_{vol}^{(1)} = k_1/\eta_{vol}$, $Q_{mirr}^{(1)} = k_1/\eta_{mirr}$, $k_1 = \Omega_1/v_h$, and $Q_{magn}^{(1)}$ is the negative (at $K > 0$) magnetic "quality factor" of this mode (for which phasing starts first).

This quality factor is given by

$$Q_1 = -k_1/\alpha_1(P, \mathbf{H}, \Omega_1) \equiv -k_1[K(P, \mathbf{H})\sigma(\mathbf{H}, \Omega_1)]^{-1}, \qquad (2)$$

where α_1 is the positive (at $K(P) > 0$) quantum amplification coefficient of hypersound for the mode under consideration of hypersound and σ is the paramagnetic absorption at $P = 0$.

The expression for σ has the form [25, p. 283]

$$\sigma_{mn} = \frac{2\pi^2 C_a v^2 g(v)|\Phi_{mn}|^2}{(2S+1)\rho' v_h^3 k_B T}, \qquad (3)$$

where $v = \Omega/2\pi$, $g(v)$ is the form factor of the APR line, ρ' is the crystal density, k_B is the Boltzmann constant, and Φ_{mn} is the factor that couples an $E_m \rightleftharpoons E_n$ spin transition with a hypersonic wave of given propagation direction and polarization.

The form factor of the APR line is normalized to unity,

$$\int_0^\infty g(v)dv = 1, \qquad (4)$$

and the matrix element Φ_{mn} for an arbitrary hypersonic wave traveling along the axis $\mathbb{O}_3$ (the axis $\mathbb{O}_3$ is aligned with the z coordinate axis) of the ruby is given by

$$\Phi_{mn} = \frac{\partial}{\partial\varepsilon_{zz}}\langle\psi_m|\hat{\mathcal{H}}|\psi_n\rangle$$
$$= \frac{G_{33}}{2}(3\langle\psi_m|\hat{S}_z^2|\psi_n\rangle - S(S+1)\langle\psi_m|\psi_n\rangle), \qquad (5)$$

where ε_{zz} is the component of the elastic strain tensor, $\hat{\mathcal{H}}$ is the Hamiltonian of spin–phonon interaction [6, 9], G_{33} is the component of the Voigt spin–phonon interaction tensor [6], and $\hat{S}_z$ is the projection of the vector spin operator on the z axis.

To estimate Φ_{mn}, we use the value $G_{33} = 5.8$ cm^{-1} = 1.16×10^{-15} erg, found experimentally, and the wave functions for the $E_3 \rightleftharpoons E_2$ transition in a Cr^{3+} ion exposed to the trigonal crystal field of ruby that were calculated in [24] from ESR data. With $H = 3.92$ kOe and $\mathbf{H}$ directed at an angle $\vartheta = \vartheta_{symm}$ to the z axis, where $\vartheta_{symm} = \arccos(1/\sqrt{3}) = 54°44'$, we find from (5) that $\Phi_{mn} \sim 10^{-15}$ erg. The choice $\vartheta = \vartheta_{symm}$ refers to the so-called symmetric (or push–pull) pumping conditions [24]. Such conditions are set up owing to the equality $E_4 - E_2 = E_3 - E_1$ (Fig. 1), which takes place at $\vartheta = \vartheta_{symm}$ and provides the most efficient inversion at the transition $E_3 \rightleftharpoons E_2$ in the spin system. Eventually, with $v_S = 9.1$ GHz, $g(v_S) = 10^{-8}$ s, $C_a = 1.3 \times 10^{19}$ cm^{-3}, $\rho' = 4$ g/cm^3, $v_h = 1.1 \times 10^6$ cm/s, and $T = 1.8$ K, we find from (3) that $\sigma \approx 0.04$ cm^{-1}.

The loaded acoustic quality factor $Q_C^{(0)}$ of the ruby FPAR (with the piezoelectric film) was measured by the pulsed echo method at frequencies $\Omega = 9.0$–9.2 GHz. With $\mathbf{H} = 0$ and $P = 0$, $Q_C^{(0)}$ was found to be $(5.2 \pm 0.4) \times 10^5$ for all longitudinal acoustic modes falling into this frequency interval. Hence, $\eta \equiv \eta_{vol} + \eta_{mirr} = \Omega/Q_C^{(0)} v_h \approx 0.1$ cm^{-1}.

The parameters σ, η, and $\alpha = \alpha_{ph}$ (where α_{ph} is the value of α at which phasing starts) are obviously related as

$$\alpha_{ph} = \eta = K_{ph}\sigma, \qquad (6)$$

where K_{ph} is the critical value of the inversion ratio K for the transition $E_3 \rightleftharpoons E_2$.

Substituting $\sigma = 0.04$ cm^{-1} and $\eta = 0.1$ cm^{-1} into (6) yields $K_{ph} \approx 2.5$. This value is readily attained in the case of push–pull pumping, which provides $K_{max} \approx 3.3$ under the conditions of our experiments.

1.3. Free Phasing Regime

Since the frequency width Γ_{32} of the APR line at the spin transition $E_3 \rightleftarrows E_2$ is ≈ 100 MHz and the mode separation is as small as 300 kHz, single-mode SE changes to multimode emission even if the pump threshold is exceeded slightly. For $\Omega_P = \Omega_{CP}^{(0)} = 23.0$ GHz and $H = H_0 = 3.92$ kOe, free multimode phasing is observed even at $P \geq 50$ μW. If $\Delta_H \equiv H - H_0 \neq 0$, for the condition $K > K_{ph}$ to be fulfilled, the pump intensity (the pump source power) must be much higher (by one to two orders of magnitude).

With pumping switched on stepwise, the free phasing conditions are set in the oscillatory regime. For our system, the frequency ω_R of these damped oscillations (the so-called relaxation frequency [26]) lies in the low-frequency range, $\omega_R \approx 130$ Hz at $H = H_0$ [27–29].

In a free-running multimode phaser, the number of modes does not exceed thirty even if $P = P^{(\max)} \gg P_{ph}$. That is, the maximal width of the phasing spectrum $(30 \times 310$ kHz ≈ 10 MHz) is one order of magnitude smaller than Γ_{32}, which is explained by the well-known Tang–Statz–deMars mechanism (exhaustion of power supplies for competing modes) [26]. If magnetic field offsets are absent, free phasing proceeds under near-steady-state conditions (the integral intensity J_Σ of multimode SE, which is measured by the hypersonic converter on one of the FPAR mirrors, is virtually time-independent).

If the offset is small, $|\Delta_H| \leq 3$ Oe, the value of J_Σ also remains time-independent. With $|\Delta_H|$ increasing to ≈ 30 Oe, the integral intensity J_Σ of phonon SE oscillates weakly because few SE modes retune or decay [18]. At $|\Delta_H| \geq 30$ Oe, some of the free phasing modes in phonon microwave spectra (PMS) split.

The splits are usually equal to several kilohertz (much more rarely several tens of kilohertz), which is much less than $\Delta\Omega_N = 300$ kHz; the number of split modes is one or two (three at most) even if $|\Delta_H| = 100$ Oe; and the intensity of such modes is much lower than that of the unsplit ones. This fine structure of SE modes under the free phasing conditions is of oscillating character, with not only the spectral component amplitudes but also their frequency positions varying smoothly (the latter within 10 kHz). The only exceptions are narrow gaps in sets of control parameters $\{P, \Delta_H$, etc.$\}$, where the spontaneous cascade decay of split SE modes takes place, causing small-scale phonon turbulence [18]. In most cases, however, free phasing in a ruby phaser proceeds under steady-state conditions even if the offset $|\Delta_H| = 100$ Oe.

1.4. Resonant Destabilization of Phonon Stimulated Emission under Low-Frequency Resonance

The situation changes when the pump modulation frequency lies near $\omega_m = \omega_R = 100$ Hz, where pro-

nounced nonlinear low-frequency resonance is observed [27–29]. If the depth of periodic modulation is small, the phonon SE integral intensity $J_\Sigma(t)$ oscillates synchronously with the external force period: $J_\Sigma(t) = J_\Sigma(t + \tau_m)$, where $\tau_m = 2\pi/\omega_m$. As the pump modulation depth k_m increases, the period of $J_\Sigma(t)$ doubles according to the Feigenbaum scenario $J_\Sigma(t) = J_\Sigma(t + 2^f \tau_m)$, where $f = f(k_m)$ successively takes the values $f = 1, 2, 3, \dots$, which condense ($f \longrightarrow \infty$) in the vicinity of a critical point $k_m = k_m^{(cr)}$. A further increase in the depth of modulation ($k_m > k_m^{(cr)}$) switches the phaser into the state of deterministic chaos [27, 29]. In the case of hard excitation (for example, by a pulse of hypersound injected into the FPAR from the outside), a phaser with periodically modulated pumping exhibits SE multistability (branching of periodic and/or chaotic phasing regimes, which causes hysteresis) [28]. Finally, the collision of a strange attractor with an unstable manifold that separates the upper periodic branch generates so-called crises (step changes in the domain of attraction, which are accompanied by attractor reconfiguration) [28, 29].

However, all the above phenomena were detected in [27–29] by measuring $J_\Sigma(t)$. More detailed information about phaser destabilization by a periodic force can be extracted from the microwave spectral characteristics of phonon SE. It has been found that, when the depth of modulation increases, SE modes alternate in a stepwise manner over the entire phasing spectrum (Fig. 2), as opposed to the insignificant smooth oscillation of the amplitude and frequency of the SE spectral components in the free phasing regime. In a number of cases, the intensity of the most powerful components of nonautonomous phasing exceeds the intensity of the most powerful (central) SE component of an autonomous phaser
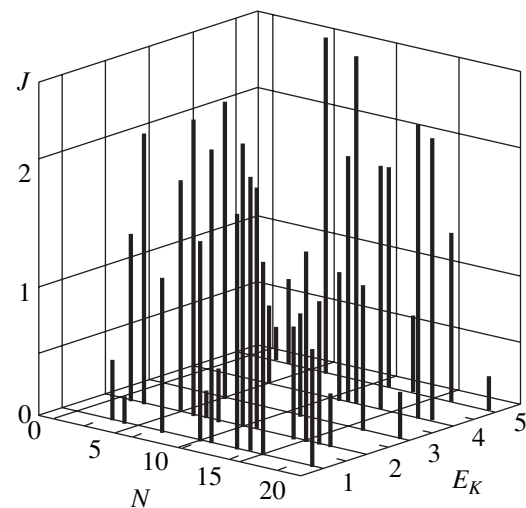

Fig. 2. Evolution of the phonon SE spectra near the low-frequency resonance at $\omega_m = 137$ Hz. The interval between sequential stages of evolution E_K is roughly equal to 1 s.

by two orders of magnitude. The evolution of the SE spectrum shown in Fig. 2 (in discrete time) illuminates the actual (complicated) pattern of motions in the spin–phonon system of an acoustic quantum oscillator, which is shaded by the integral characteristic $J_\Sigma(t)$. A similar evolution of the PMS was also obtained upon magnetic field modulation.

The mean lifetime of microwave phonon modes under the conditions shown in Fig. 2 is several tenths of a second. The modes rearrange in an irregular manner, and their distribution is not repeated. Such a chaotic evolution of the SE spectra took place over the entire frequency range $\omega_m = 70\text{–}200$ Hz, where resonant low-frequency destabilization of phasing and, accordingly, chaotic oscillations of $J_\Sigma(t)$ were previously observed [27–29].

1.5. Resonant Destabilization of Phonon Stimulated Emission in the Case of Ultralow-Frequency Resonance

Under the ultralow-frequency ($\omega_m \approx 10$ Hz) modulation of pump or magnetic field, the stabilization of phonon stimulated emission assumes another, laminar, character. Unlike the case of low-frequency resonance, resonance at ultralow frequencies is characterized by an extremely high correlation of spectral motions. If the modulation frequency ω_m is precisely tuned to the vertex ω_λ of ultralow-frequency (ULF) resonance and the depth of modulation is high (close to 100%), the phonon microwave spectra narrow roughly fourfold and contain no more than six or seven modes of phonon SE.

With a small mismatch in terms of pump modulation frequency, $\Delta_\lambda \equiv \omega_m - \omega_\lambda$, these narrow SE spectra exhibit regular reconfigurations with intriguing fea-

tures. It was found experimentally that the period $\tau_d^{(\lambda)}$ of reconfiguration varies by several orders of magnitude when Δ_λ varies by no more than 1 Hz. In addition, the period $\tau_d^{(\lambda)}$ turned out to be incommensurate with the external force period $\tau_m \equiv 2\pi/\omega_m$ (that is, the frequency $\omega_d^{(\lambda)} \equiv 2\pi/\tau_d^{(\lambda)}$ generally is not a harmonic or subharmonic of the driving force frequency ω_m). In experiments, this shows up as the instability of states that have rational ratios $\tau_d^{(\lambda)}/\tau_m$. The essence of the PMS self-reconfiguration is the periodic unidirectional displacement (with the period $\tau_d^{(\lambda)}$) of the range of microwave modes (this range typically comprises from three to seven modes) along the frequency axis if $\omega_m \approx \omega_\lambda$.

It is noteworthy that the frequency position of each of the modes remains nearly unchanged (if higher order dynamic effects due to the nonstationary fine structure of the SE spectra [23] are disregarded): only the position of the spectral part with phasing modes changes. Thus, the ignition of new FPAR modes at one edge of the PMS is accompanied by the extinguishing of the same number of modes at the opposite end of the PMS. Such a motion lasts until SE ceases completely in some range of microwave frequencies. After a relatively short period of complete absence of stimulated emission (a period of refractoriness), the process of spectrum global self-reconfiguration is repeated, starting from the same position on the frequency axis. The period $\tau_d^{(\lambda)}$ of these unidirectional spectral motions remains the same if the set of control parameters does not change. On the screen of a spectrum analyzer, this evolution of the SE spectrum appears as the periodic motion of a mode cluster. Typical sequences of SE spectra under ULF resonance conditions ($\omega_\lambda = 9.79$ Hz) at $\Delta_\lambda = -0.23$ Hz and in the absence of static magnetic field mismatch are shown in Fig. 3.

An instantaneous set of SE modes forms a cluster of a certain width, and this width varies insignificantly during the motion. At the same time, the set of SE modes that form a cluster constantly varies (Fig. 3). As follows from Fig. 3, the self-reconfiguration of the PMS is imposed on irregular oscillations of the SE mode intensity. When the sweep range of the spectrum analyzer is decreased by two or three orders of magnitude, weak irregular motions of SE modes along the frequency axis and sometimes a split of one of these modes are observed (Fig. 4). The modes split when their intensity is very low, i.e., immediately before extinguishment (because of this, the instrument noise is noticeable in Fig. 4). Obviously, such fine effects cannot be seen on the panoramic spectra in Fig. 3 (the fine structure of the PMS in a ruby phaser was studied in [18, 23]). In general, it can be said that large-scale ordered (laminar) motions of SE spectra in a phaser with ULF pump modulation are imposed on small-scale irregular processes.
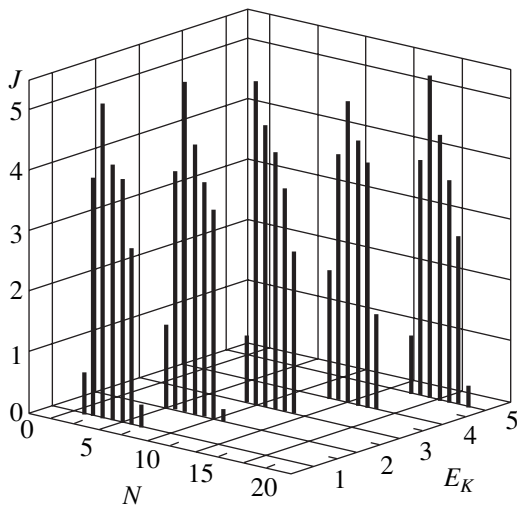

Fig. 3. The same as in Fig. 2, but near the ULF resonance at $\omega_m = 9.56$ Hz. The step between E_K is 2.5 s.

Similar large-scale laminar self-reconfigurations of phonon microwave spectra in a nonautonomous ruby phaser were observed in experiments at $\Delta_\lambda = +0.23$ Hz; however, the cluster moved in the opposite direction. Further investigations showed that the sign of the derivative $d\Omega_V/dt$ (here, Ω_V is the center frequency of a mode cluster) strictly correlates with the sign of the frequency offset of the external force from resonance: $\operatorname{sgn}[d\Omega_V/dt] = -\operatorname{sgn}\Delta_\lambda$. Approaching the exact ULF resonance discovered ($|\Delta_\lambda| \longrightarrow 0$), the self-reconfiguration period $\tau_d^{(\lambda)}$ takes giant values. Direct measurements $\omega_d^{(\lambda)} \equiv 2\pi/\tau_d^{(\lambda)}$ gave $\inf(\omega_d^{(\lambda)}) < 10^{-4}$ Hz (one self-reconfiguration period is ≈ 3 h). All the experiments were performed with superfluid helium ($T = 1.8$ K) in order to remove problems associated with boiling a cryogenic liquid.

Such a character of the PMS self-reconfiguration persists over a wide range of Δ_H. Moreover, at $|\Delta_H| < 10$ Oe, even the value of ω_λ remains constant (close to 9.8 Hz). Only when the magnetic field mismatch increases further does the resonance frequency ω_λ decrease tangibly (about twofold for $|\Delta_H| \approx 60$ Oe). It is significant that the above dependence of the direction of motion of a mode cluster on the modulation frequency detuning, $\operatorname{sgn}[d\Omega_V/dt] = -\operatorname{sgn}\Delta_\lambda$, remains valid.

1.6. Harmonics of Ultralow-Frequency Resonance

Along with intermittent phasing conditions (including the periods of refractoriness), we observed conditions under which at least one phasing mode appears in the starting range before phasing in the final range disappears. In other words, this means that two narrow mode clusters whose virtual tops V_1 and V_2 move in the same direction and with the same velocity, $d\Omega_{V1}/dt = d\Omega_{V2}/dt$, are present on the frequency axis simultaneously. The same effect was observed for the first three harmonics of ULF resonance: at $\omega_m \approx \omega_{2s\lambda} \equiv 2s\omega_\lambda$, where $s \in \{1, 2, 3\}$ (the lowest $2s\lambda$ resonances). As for the fundamental λ resonance ($\omega_m = \omega_\lambda$), the corresponding periods of PMS reconfiguration, $\tau_d^{(2s\lambda)}$, in the case of our phaser are incommensurate with the external force period $\tau_m \equiv 2\pi/\omega_m$ and increase to 100 s or more if the detuning absolute value $|\Delta_L^{(2s\lambda)}| \equiv |\omega_m - \omega_{2s\lambda}|$ is small (less than 0.05–0.10 Hz). In this case, too, the sign of $d\Omega_V/dt$ (or $d\Omega_{V1}/dt$ and $d\Omega_{V2}/dt$) is opposite to the sign of $\Delta_L^{(2s\lambda)}$.

With $4 \leq s \leq 11$, the driving force frequency ω_m falls into the range of very wide relaxation resonance mentioned above (which prevails at frequencies between 70 and 200 Hz). However, for $s > 11$, i.e., when the driving
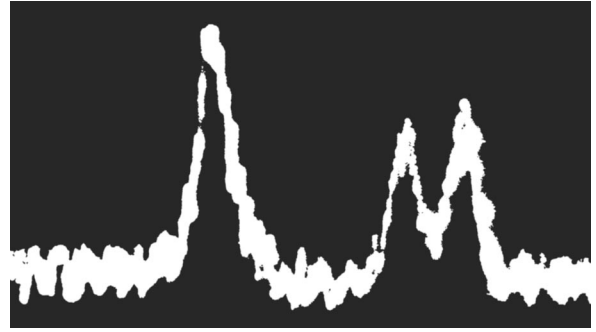

Fig. 4. Fine structure of the phonon SE spectrum. The sweep along the abscissa axis is 10 kHz.

force leaves this range of resonant destabilization, our experiments again distinctly revealed narrow-band resonant responses of the phasing system to the external effect not only at frequencies $\omega_{2s\lambda}$ but also at odd harmonics $\omega_{(2s+1)\lambda} \equiv (2s+1)\omega_\lambda$. These responses somewhat differ from those in the case of the fundamental λ resonance and its first even harmonics. For example, at $s > 11$, the deviation of Ω_V is, as a rule, no greater than one or two mode separations (i.e., no greater than 0.3–0.6 MHz) and phasing modes experience deep periodic modulation (of depth 50% or more).

At $s > 11$ and a detuning of ≈ 1 Hz, this self-modulation is fast (its period is about a second). As ω_m approaches the top of each of the $2s\lambda$ or $(2s+1)\lambda$ resonances (at the same $s > 11$), the self-modulation period τ_{sm}, as well as the period $\tau_d^{(\lambda)}$ of mode alternation for the fundamental frequency and lower even harmonics ($s < 4$), increases monotonically. The highest values of τ_{sm}, which were reliably observed at $11 < s < 20$, reached several minutes. Note that the same intermittent conditions of phonon stimulated emission (but with smaller self-modulation depths) were also found for the first two odd harmonics ($\omega_m \approx 3\omega_\lambda$ and $\omega_m \approx 5\omega_\lambda$), where "intense" self-reconfigurations are absent.

2. RESULTS AND DISCUSSION

Our experimental data for ULF modulation suggest self-organization in the spin–phonon system of a phaser in terms of energy exchange between modes. It should be emphasized that highly organized collective motions in the spin–phonon system are observed not only for each of the microwave acoustic modes (as in the case of steady-state multimode phasing) but also at the global level, where all SE modes obey the same rhythm, the frequency of which is not a harmonic (subharmonic) of an external perturbation. In other words, if upon normal steady-state multimode phasing in an autonomous phaser there exist N virtually independent microwave oscillators (each corresponding to a specific SE mode), upon a resonant ULF perturbation of the pump or mag-

netic field, these oscillators behave consistently (cooperatively).

The features of collective motions in a phaser suggest that this effect is similar to antiphase dynamics processes [30–32], which were discovered previously in multimode lasers. In the simple case of double-mode lasing [30], antiphase dynamics appears as consistent oscillations of modes strictly in antiphase (hence, the name of the effect). In N-mode systems, antiphase motions may be much more complicated (see, e.g., [31, 32]); however, the general nontrivial tendency, namely, coherent unidirectional SE mode oscillations with a time delay $\tau_d^{(\lambda)}/N$ between nearest neighbors, still persists.

As was found in this work, energy exchange between modes in a phaser results in an additional characteristic frequency ω_λ, which is much lower than the relaxation frequency ω_R. Collective motions are excited when the parameters of the phasing system are modulated at frequencies close to ω_λ. The same is true for lasers exhibiting antiphase dynamics [30–32]. Accordingly, phonon spectral self-reconfigurations may be treated as the occurrence of antiphase states of stimulated emission in the FPAR when the spatial distribution of stationary modes is destabilized by an external force at $\omega_m \approx \omega_\lambda$. Moreover, the value of ω_λ estimated by formulas given in [30] is one order of magnitude higher than ω_R, which is also in agreement with our experimental data for nonautonomous phasing.

It should be noted, however, that the laser model of the nonlinear dynamics of phonon microwave stimulated emission cannot describe adequately all features of self-organization in a phaser near ULF resonance, although it gives a satisfactory estimate of ω_λ and predicts more or less accurately the character of mode motions. For a better understanding of self-organization in a ruby laser, one should consider the unusual hierarchy of spin reservoirs [33], which are responsible for the specific saturation of quantum transitions in the microwave range. Essentially, all nonlinearities appearing in microwave resonant interactions of the signal acoustic field and the electromagnetic pump field with the electron Zeeman reservoir $\tilde{Z}_E$ [33] of Cr^{3+} ions in ruby "sense" to one extent or another the presence of slowly relaxing Al^{27} nuclei, as revealed in earlier experiments on hypersound quantum amplification [10, 11, 17].

The reason for this sensitivity is direct thermal contact [33] between the nuclear Zeeman reservoir $\tilde{Z}_N$ and the dipole–dipole interaction reservoir $\tilde{D}_E$, owing to which the energy exchange $\tilde{Z}_E \Longleftrightarrow \tilde{D}_E \Longleftrightarrow \tilde{Z}_N$ takes place. It is important that the heat capacities of the reservoirs $\tilde{Z}_N$ and $\tilde{Z}_E$ are comparable to each other, although the frequency of nuclear magnetic resonance (≈ 10 MHz) for Al^{27} is three orders of magnitude lower

than the ESR and APR frequencies for Cr^{3+} ions at $H \approx 4$ kOe. This is because the concentration of impurity paramagnetic Cr^{3+} ions in pink ruby is as low as several hundredths of a percent; that is, for one electron spin, there are several thousands of nuclear spins. As a result, the inertial nuclear system, while unseen in direct ESR and APR measurements, participates in all population redistributions over electron spin levels (in more exact terms, over a quasi-continuous set of sublevels due to dipole–dipole interactions [33]).

Turning back to ULF resonance, we may assume that the decrease in its frequency at large mismatches Δ_H is observed when nuclear degrees of freedom are involved in energy exchange between modes, which goes through $\tilde{Z}_E$ and $\tilde{D}_E$ and causes self-organization in the phasing medium. In fact, the most important feature of interaction between the electron and nuclear subsystems of ruby is the strong dependence of the spin temperatures (in all the three reservoirs) on detuning in the phaser [10, 11, 17]. It is also noteworthy that the combined effect of the low-energy reservoirs $\tilde{D}_E$ and $\tilde{Z}_N$ on phonon SE in a phaser differs radically from the effect of similar reservoirs in optical lasers (see, e.g., [34]). The point is that, in our system, the heat capacity and inertia of the low-energy reservoirs differ from those of the high-energy reservoir $\tilde{Z}_E$ much more than in CO_2 lasers [34]. First, in our system, as was noted above, for one active center Cr^{3+} there are several thousands of magnetic nuclei Al^{27}, which make a low-energy reservoir "heavier," while in CO_2 lasers the low- and high-energy reservoirs are formed by the same molecules (in CO_2 lasers, the high-energy reservoir is formed by rotational–vibrational degrees of freedom of CO_2 molecules). To put it otherwise, a laser has analogues to the reservoirs $\tilde{Z}_E$ and $\tilde{D}_E$ but does not have an analogue to the reservoir $\tilde{Z}_N$. Second (and most important), the relaxation time of the low-energy reservoirs is much longer (because of the nuclei) than that of the high-energy one, while for CO_2 lasers [34] the reverse is true. Therefore, in a ruby phaser, the combined inertial reservoir $\tilde{D}_E + \tilde{Z}_N$ can be involved in governing self-organization processes, which the slowest processes always are.

CONCLUSIONS

We experimentally studied the influence of an external periodic force on the dynamics of phonon microwave stimulated emission in a pink ruby phaser. It is shown that the periodic modulation of pump at a frequency $\omega_m = 70$–200 Hz randomizes energy exchange between modes in the spin–phonon system because of resonant destabilization of the phaser near its relaxation resonance frequency ω_R. In this range of destabiliza-

tion, the width of the phonon microwave spectrum does not change. For ULF pump modulation ($\omega_m \approx \omega_\lambda \approx$ 10 Hz), a qualitatively new type of phonon SE destabilization is discovered. First, the PMS narrows considerably (almost four times). Fast chaotic switchings of phasing modes at $\omega_m \approx \omega_R$ near the ULF resonance give way to self-organization in intermode interaction (unlike conventional "intramode" self-organization, which is typical of multimode lasers). Self-organization appears as consistent regular oscillations of each of the SE modes with a time delay $\tau_d^{(\lambda)}/N$. This appears as the motion of a mode cluster in the spectral space. The total self-reconfiguration cycle depends considerably on $\omega_m - \omega_\lambda$ and changes by several orders of magnitude when $|\omega_m - \omega_\lambda|$ changes by several percent. The same processes were observed for the first three even harmonics of the ULF resonance. For higher even harmonics, as well as for all odd harmonics, intermittent periodic modulation is found. Its period may exceed the period of the driving force by several orders of magnitude. The results obtained are treated in terms of the antiphase dynamics of phonon SE. The effect of the Al^{27} magnetic nuclear subsystem in the ruby crystalline matrix on the reconfiguration of the phasing spectrum in a nonautonomous phaser (in particular, on the dependence of ω_λ on the magnetic field) is discussed.

ACKNOWLEDGMENTS

REFERENCES

1. C. H. Townes, in *Proceedings of the 1st International Symposium on Quantum Electronics, New York, 1960*, p. 405.
2. C. Kittel, Phys. Rev. Lett. **6**, 449 (1961).
3. Yu. E. Lozovik, Usp. Fiz. Nauk **71**, 1373 (2001) [Phys. Usp. **44**, 1307 (2001)].
4. E. B. Tucker, Phys. Rev. Lett. **6**, 547 (1961).
5. N. S. Shiren, Appl. Phys. Lett. **7** (5), 142 (1965).
6. E. Taker, *Physical Acoustics*, Ed. by W. P. Mason (Academic, New York, 1968; Mir, Moscow, 1969), Vol. 4A, pp. 63–138.
7. P. D. Peterson and E. H. Jacobsen, Science **164**, 1065 (1969).
8. E. M. Ganapol'skiĭ and D. N. Makovetskii, Dokl. Akad. Nauk SSSR **217**, 303 (1974) [Sov. Phys. Dokl. **19**, 433 (1974)].
9. V. A. Golenishchev-Kutuzov, V. V. Samartsev, N. K. Solovarov, and B. M. Khabibulin, *Magnetic Quantum Acoustics* (Nauka, Moscow, 1977).
10. E. M. Ganapol'skiĭ and D. N. Makovetskii, Fiz. Tverd. Tela (Leningrad) **24**, 1960 (1982) [Sov. Phys. Solid State **24**, 1119 (1982)].
11. D. N. Makovetskii, Candidate's Dissertation (Physicotechnical Inst. of Low Temp., Acad. Sci. of Ukr. SSR, Kharkov, 1984).
12. D. N. Makovetskii and K. V. Vorsul', Zh. Tekh. Fiz. **61** (1), 86 (1991) [Sov. Phys. Tech. Phys. **36**, 50 (1991)].
13. D. N. Makovetskii, A. A. Lavrinovich, and N. T. Cherpak, Zh. Tekh. Fiz. **69** (5), 101 (1999) [Tech. Phys. **44**, 570 (1999)].
14. E. B. Tucker, in *Proceedings of the 3rd International Congress on Quantum Electronics*, Ed. by P. Grivet and N. Bloembergen (Dunod, Paris, 1964), Vol. 2, pp. 1787–1800.
15. E. M. Ganapolskii and D. N. Makovetskii, Solid State Commun. **15**, 1249 (1974).
16. B. Fain, Phys. Rev. B **26**, 5932 (1982).
17. E. M. Ganapol'skiĭ and D. N. Makovetskii, Zh. Éksp. Teor. Fiz. **72**, 203 (1977) [Sov. Phys. JETP **45**, 106 (1977)].
18. D. N. Makovetskii, in *Proceedings of the 4th International Symposium on Physics and Engineering of Millimeter and Submillimeter Waves (MSMW-2001), Kharkov, 2001*, pp. 762–764.
19. W. Horsthemke and R. Lefever, *Noise Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology* (Springer-Verlag, Heidelberg, 1984; Mir, Moscow, 1987).
20. L. M. Pecora and T. L. Carrol, Phys. Rev. Lett. **67**, 945 (1991).
21. D. N. Makovetskii, Ukr. Fiz. Zh. **43**, 537 (1998).
22. D. N. Makovetskii, Pis'ma Zh. Tekh. Fiz. **27** (12), 57 (2001) [Tech. Phys. Lett. **27**, 511 (2001)].
23. D. N. Makovetskii, Radiofiz. Élektron. **4** (2), 92 (1999).
24. A. E. Siegmann, *Microwave Solid-State Masers* (McGraw-Hill, New York, 1964; Mir, Moscow, 1966).
25. S. A. Al'tshuler and B. M. Kozyrev, *Electron Paramagnetic Resonance* (Fizmatgiz, Moscow, 1961; Academic, New York, 1964).
26. C. L. Tang, H. Statz, and G. deMars, J. Appl. Phys. **34**, 2289 (1963).
27. E. M. Ganapol'skiĭ and D. N. Makovetskii, Zh. Tekh. Fiz. **59** (10), 202 (1989) [Sov. Phys. Tech. Phys. **34**, 1220 (1989)].
28. E. M. Ganapol'skiĭ and D. N. Makovetskii, Zh. Tekh. Fiz. **62**, 187 (1992) [Sov. Phys. Tech. Phys. **37**, 218 (1992)].
29. D. N. Makovetskii, Radiofiz. Élektron. **6** (1), 124 (2001).
30. M. Georgiou, P. Mandel, and K. Otsuka, IEEE J. Quantum Electron. **30**, 854 (1994).
31. B. A. Nguyen and P. Mandel, Opt. Commun. **138**, 81 (1997).
32. A. G. Vladimirov, E. A. Viktorov, and P. Mandel, Phys. Rev. E **60**, 1616 (1999).
33. V. A. Atsarkin and M. I. Rodak, *Problems of Magnetic Resonance* (Nauka, Moscow, 1978), pp. 187–205.
34. F. T. Arecchi, W. Gadomski, R. Meucci, and J. A. Roversi, Opt. Commun. **70**, 155 (1989).

Translated by V. Isaakyan

RADIOPHYSICS

Attenuation of Electromagnetic Waves in a Semiconductor Superlattice in a Magnetic Field

O. V. Shramkova

Usikov Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine,
Kharkov, 61085 Ukraine
e-mail: oksana@ire.kharkov.ua
Received June 19, 2003

Abstract—Magnetoplasma waves and inhomogeneous (complex) waves in an infinite semiconductor superlattice in a magnetic field are studied, and dispersion curves are obtained. It is shown analytically and numerically that, in periodic structures, among inhomogeneous waves, there are many complex waves for which the imaginary part of the wave vector is greater than the real part. The effect of dissipation in a medium on the dispersion curves of magnetoplasma waves is examined. The dependence of the minimum phase velocity on the collision frequency and the magnetic field strength is studied. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

For many years, periodic layered media have attracted great interest from researchers. These structures belong to a new type of artificial materials whose physical characteristics cannot be achieved in natural semiconductors, because their properties depend on both the physical parameters of the materials of which they are fabricated and on the layer dimensions and structure period. The specific features of periodic layered structures are attributed to their translation symmetry. Such structures are widely used in modern millimeter- and submillimeter-wave devices, antennas, optics and optoelectronics, and X-ray technique.

We consider waves in a periodic semiconductor superlattice in a magnetic field. Magnetoplasma waves in superlattices have been studied in many works. For arbitrary orientations of the lattice, applied magnetic field, and propagation direction, the problem still remains unresolved. Electromagnetic properties of superlattices with conductive layers in a magnetic field were studied in [1, 2]. In those papers, the magnetic field directions coincided with the direction of the structure periodicity and the wave propagation direction. For this geometry, analytical expressions can only be obtained for waves whose wave vector is parallel or perpendicular to the layers. It was shown in [3] that, when the magnetic field is perpendicular to the periodicity direction and a wave propagates in a plane perpendicular to the magnetic field, a specific band structure of the spectrum is observed. Two regions of bulk magnetic polaritons were also described. These results coincide with those for bulk polaritons described in later papers [4, 5]; however, the frequency positions of the passbands were not analyzed. The properties of such structures with a magnetic field vector parallel to the layers are studied in [6]. The coefficient of reflec-

tion from a semi-infinite semiconductor structure consisting of semiconductor and dielectric layers in an external magnetic field perpendicular to the direction of structure periodicity was studied in [7]. However, all the works mentioned above ignore dissipation, which affects the wave dispersion. At the same time, it was shown [8–10] that dissipation limits the maximum value of the wavenumber, i.e., the minimum phase velocity $v_{\mathrm{ph}} = \omega/k_x$.

In the next section, we will study inhomogeneous (complex) magnetoplasmons in a semiconductor superlattice. Let us recall what inhomogeneous plane waves are. In study of the propagation of plane electromagnetic waves in infinite media, the dependence of the field on coordinates is given by the factor $e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t}$, where $\mathbf{k}$ is the complex wave vector,

$$\mathbf{k} = \mathbf{k}' + i\mathbf{k}'',$$

and $\mathbf{k}'$ and $\mathbf{k}''$ are the real vectors [11].

From Maxwell's equations, we have

$$k^2 = \frac{\omega^2}{c^2}\varepsilon. \tag{1}$$

For real permittivities, this expression is meaningful if $\mathbf{k}' \cdot \mathbf{k}'' = 0$. Waves for which planes of constant phase (i.e., the planes perpendicular to the vector $\mathbf{k}'$) and planes of constant amplitude (i.e., the planes perpendicular to the vector $\mathbf{k}''$) are orthogonal are referred to as inhomogeneous waves. In the literature, these waves are called complex waves [12].

Inhomogeneous plane waves were addressed in [12, 13]. The properties of inhomogeneous plane waves in a periodic structure consisting of dielectric layers and semiconductor layers whose permittivity is frequency-dependent and may become negative were first considered in [14]. The permittivity is negative at fre-

quencies below the plasma frequency. It can be seen from Eq. (1) that $\mathbf{k}'' > \mathbf{k}'$ at $\varepsilon < 0$; i.e., the imaginary part of the wave vector is greater than its real part. This circumstance imparts a number of interesting features to the propagation of complex waves.

This paper differs from those mentioned above in that it considers an infinite periodic semiconductor structure in a magnetic field. It is assumed that the magnetic field is parallel to the layers. We calculate the dispersion characteristics of complex magnetoplasma waves. The effect of attenuation on the dispersion properties of inhomogeneous magnetoplasmons is studied.

DISPERSION RELATION

We will consider an infinite periodic structure whose periodic cell consists of a semiconductor layer of thickness d_1 and a dielectric layer of thickness d_2. Let the structure be exposed to an external magnetic field parallel to the y axis. The z axis is perpendicular to the layer boundaries. The magnetoplasma waves propagate in the xz plane [15]. Propagation of electromagnetic waves in such a structure is described by Maxwell's equations written for each layer and the condition that tangential components of the electric and magnetic fields be continuous on all the boundaries of the structure. We will seek a solution to this system of equations in the form $\exp(ik_x x + ik_{z1,2} z - i\omega t)$. We assume that the structure is uniform in the y direction; hence, $\partial/\partial y = 0$. Then, Maxwell's equations split into independent equations for two modes with different polarizations. We consider the mode with the components E_x, E_z, and H_y. The permittivity tensor of the semiconductor layer can be written as [16]

$$\varepsilon_{xx} = \varepsilon_{zz} = \varepsilon_{\parallel} = \varepsilon_0\left[1 - \frac{\omega_p^2(\omega + i\nu)}{\omega[(\omega + i\nu)^2 - \omega_H^2]}\right],$$

$$\varepsilon_{xz} = -\varepsilon_{zx} = \varepsilon_{\perp} = -i\varepsilon_0\frac{\omega_p^2\omega_H}{\omega[(\omega + i\nu)^2 - \omega_H^2]},$$

where ε_0 is the part of the permittivity attributed to the lattice, ω_p is the plasma frequency, ω_H is the cyclotron frequency, and ν is the collision frequency.

The permittivity ε_2 of the second layer is constant. To satisfy the boundary conditions, we use the transfer matrix method (this matrix relates fields at the beginning and end of the period) and Floquet's theorem, which allows for the periodicity. The dispersion relation for an infinite medium in a magnetic field has the form [15]

$$\cos \bar{k}d = \cos k_{z1}d_1 \cos k_{z2}d_2 - \frac{\varepsilon_{f1}\varepsilon_2}{2k_{z1}k_{z2}}\left[\left(\frac{k_{z1}}{\varepsilon_{f1}}\right)^2 + \left(\frac{k_{z2}}{\varepsilon_2}\right)^2\right]$$

$$- k_x^2\left(\frac{\varepsilon_{\perp 1}}{\varepsilon_{\parallel 1}\varepsilon_{f1}}\right)^2\right]\sin k_{z1}d_1 \sin k_{z2}d_2,$$

(2)

where index 1 refers to the semiconductor layers and index 2, to the dielectric layers;

$$k_{z1} = \left[\frac{\omega^2\varepsilon_{V1}}{c^2} - k_x^2\right]^{1/2}, \quad k_{z2} = \left[\frac{\omega^2\varepsilon_2}{c^2} - k_x^2\right]^{1/2},$$

$\varepsilon_{V1} = \varepsilon_{\parallel 1} + \varepsilon_{\perp 1}^2/\varepsilon_{\parallel 1}$ is the Voigt permittivity; k_{z1} and k_{z2} are the transverse wave numbers of the first and second layer, respectively; $d = d_1 + d_2$ is the structure period; and k is the wavenumber averaged over the period.

The analysis of expression (2) reveals the characteristic frequencies [3]

$$\omega_{01,02} = \mp\frac{\omega_H}{2} + \sqrt{\frac{\omega_H^2}{4} + \omega_{ps}^2},$$

(3)

$$\omega_{\infty} = \sqrt{\omega_p^2 + \omega_H^2 + \nu^2},$$

which are the limiting frequencies for oscillations in the superlattice (they determine the asymptotes $\omega = \omega_{01}$, $\omega = \omega_{\infty}$, and $\omega = \omega_{02}$ of the dispersion curves). Here, $\omega_{ps} = \omega_p(\varepsilon_{01}/(\varepsilon_{01} + \varepsilon_2))^{1/2}$ is the frequency of the surface plasmon on the semiconductor–dielectric boundary. The behavior of the dispersion curves is determined by the magnetic field strength. For $\omega_H < \omega_{cr}$, we have $\omega_{02} < \omega_{\infty}$, and, for $\omega_H > \omega_{cr}$, we have $\omega_{02} > \omega_{\infty}$. The critical frequency

$$\omega_{cr} = \omega_p\frac{\varepsilon_2}{\varepsilon_{01}}\left(1 - \frac{\varepsilon_2}{\varepsilon_{01}}\right)^{-1/2}$$

is determined from the condition $\omega_{02} = \omega_{\infty}$.

Figure 1 shows dispersion curves for magnetoplasma waves with allowance for the delay and without allowance for attenuation in the medium. The calculations were performed for the following parameters: The first layer was an InSb semiconductor with $\varepsilon_{01} = 17.8$,
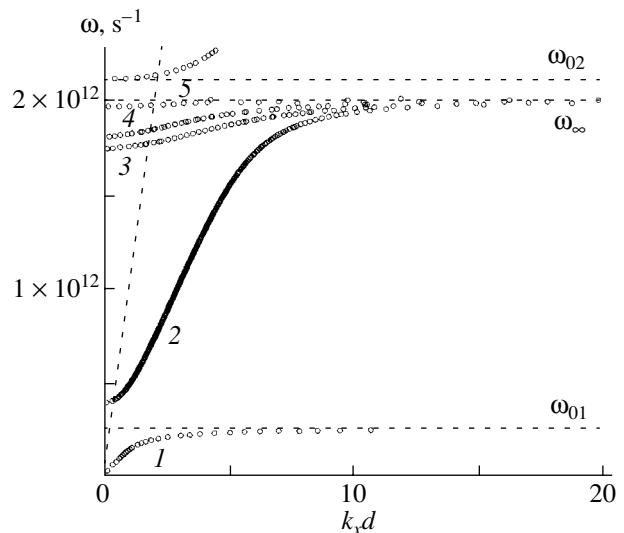

Fig. 1. Dispersion curves for magnetoplasma waves.

$\omega_{p1} = 10^{12}$ s^{-1}, and $d_1 = 0.015$ cm. The second layer was a dielectric with $\varepsilon_2 = 2$ and $d_2 = 0.005$ cm. The magnetic field was $H_0 = 1000$ Oe ($\omega_H > \omega_{cr}$). It was also assumed that $kd = 0$. At frequencies $\omega < \omega_{01}$, the fields of the propagating waves decrease exponentially with the distance from the layer boundaries. These waves are collective surface magnetoplasmons, whose dispersion curves tend to the asymptote $\omega = \omega_{01}$ (curve 1). In the frequency band $\omega_{01} < \omega < \omega_{02}$, two characteristic regions can be distinguished [7]. These regions are divided by the line $k_{z2} = 0$, corresponding to a wave propagating with the speed of light (the oblique dashed line in Fig. 1). In the first region, the waves propagate in the semiconductor and dielectric layers as in a waveguide ($k_{z1,2}^2 > 0$) and allowed and forbidden zones are formed due to the geometric resonance conditions, under which the layer thickness is a multiple of half-waves. In the second region, $k_{z1}^2 > 0$, while $k_{z2}^2 < 0$, and the fields of the surface polaritons tunnel through the dielectric layer. The frequency of magnetoplasma waves is seen to approach the hybrid frequency ω_∞ (curves 2–5) as $k_x d$ increases. This property of the magnetoplasmon spectrum is described in [17].

ALLOWANCE FOR DISSIPATION

Allowance for dissipation in a medium changes the behavior of the dispersion curves. Figure 2 shows the dispersion curves with allowance for the collision frequency ($\nu = 10^{11}$ s^{-1}). It is assumed that $k_x = k_x' + i k_x''$ (k_x'' is related to dissipation). The solid lines show the real part of the wavenumber, $k_x' d$, as a function of frequency; the dashed lines refer to the imaginary part of the wavenumber, $k_x'' d$. It follows from Fig. 2 that, at $k_x' d \leq 1$, the imaginary part can be greater than the real part k_x'. As the frequency and k_x' increase, k_x'' almost vanishes. In particular, at the point E, the relative attenuation is $k_x''/k_x' \approx 0.02$. As the frequency approaches the hybrid frequency, the attenuation increases (at the point F, $k_x''/k_x' \approx 0.32$), while k_x'' differs little from k_x'. Thus, the relative wave attenuation k_x''/k_x' is small and is almost equal to unity at low and high frequencies.

These plots differ from those for a nondissipative medium in that they have a turning point at high $k_x' d$, where k_x' takes a maximum value. To estimate $(k_x)_{max}$, we assume that $\omega \approx \omega_\infty$ and $k_x' \gg (\omega_\infty/c)\sqrt{\varepsilon_{V1,2}}$. Then, we obtain

$$k_{z1} = k_{z2} \approx i k_x \approx i(k_x' + i k_x'') \tag{4}$$

and dispersion relation (2) takes the form

$$\cos \bar{k} d = \cosh k_x d_1 \cosh k_x d_2 + \frac{\varepsilon_{V1}\varepsilon_2}{2}$$
$$\times \left[\frac{1}{\varepsilon_{V1}^2} + \frac{1}{\varepsilon_2^2} + \left(\frac{\varepsilon_{\perp 1}}{\varepsilon_{\|1}\varepsilon_{V1}} \right)^2 \right] \sinh k_x d_1 \sinh k_x d_2. \tag{5}$$

At high k_x', we have $\cosh k_x' d_{1,2} \approx \sinh k_x' d_{1,2} \approx \exp(k_x' d_{1,2})/2$. Then,

$$k_x' = -\frac{1}{d} \ln \frac{2\varepsilon_{01}(2\omega_\infty^2 - \omega_p^2)\nu}{A},$$

$$A = \sqrt{\frac{\omega_p^4}{\omega_\infty^2}\left((1 + \varepsilon_2)\omega_\infty^2 - \frac{\varepsilon_{01}^2}{2}(\omega_\infty^2 - \omega_p^2) \right)^2 + 4\nu^2\left((1 + \varepsilon_2)\omega_\infty^2 + \frac{\varepsilon_{01}^2}{2}(\omega_\infty^2 - \omega_p^2) \right)^2}, \tag{6}$$

$$k_x'' = \frac{1}{d}\arctan \frac{\omega_p^2}{\omega_\infty^2}\frac{1}{2\nu}\frac{(1 + \varepsilon_2)\omega_\infty^2 - \frac{\varepsilon_{01}^2}{2}(\omega_\infty^2 - \omega_p^2)}{(1 + \varepsilon_2)\omega_\infty^2 + \frac{\varepsilon_{01}^2}{2}(\omega_\infty^2 - \omega_p^2)}.$$

Here, $|\cos kd| \approx 1$. From formulas (6), we obtain that, as $\nu_1 \longrightarrow 0$, $k_x' d \longrightarrow \infty$ and $k_x'' d \longrightarrow \pi/2$.

The analysis of formulas (6) allows us to determine the physical meaning of the turning point on the dispersion curves. The fact is that k_x'' now determines the phase of the field along the z axis: $\exp(k_x' - i k_x'')z$. Therefore, when $\nu \neq 0$, the phase increment across the semiconductor layer is $\approx \pi/2$. This means that the wave that passes through the semiconductor layer and reflects from its boundary is in antiphase with the wave incident on the layer; i.e., the waves cancel each other. This effect creates a forbidden band for $k_x' > k_{x max}'$, and a turning point appears on the dispersion curves.

The behavior of the dispersion curves is determined by the magnetic field strength H_0. Therefore, the phase
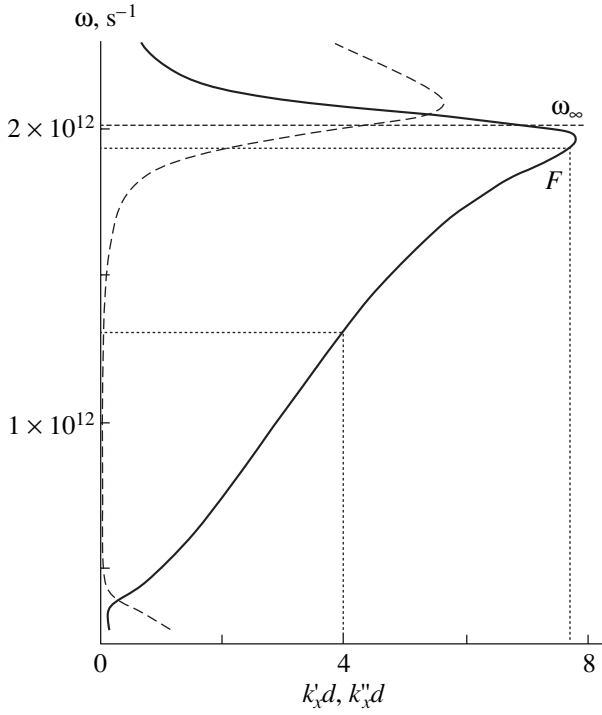
Fig. 2. Dispersion curves for magnetoplasmons with allowance for spatial attenuation: $\omega(k_x')$ (solid lines) and $\omega(k_x'')$ (dashed lines).

Fig. 3. Minimum phase velocity of magnetoplasmons vs. frequency and magnetic field.

velocity depends on the cyclotron frequency. Figure 3a shows $v_{ph\,min}$ versus the cyclotron frequency ω_H at $\nu =$ (*1*) 10^{11} and (*2*) 2×10^{11} s^{-1}. It can be seen that, when $\omega_H < \omega_{cr}$, $v_{ph\,min}$ decreases monotonically with increasing applied magnetic field. The phase velocity reaches its minimum value at $\omega_H = \omega_{cr}$. Further, the phase velocity increases with the magnetic field. To calculate the minimum phase velocity $v_{ph\,min}^*$, we use the expression for k_x' (the first of formulas (6)):

$$v_{ph\,min}^* = -\omega d / \ln \frac{2\varepsilon_{01}(2\omega_\infty^2 - \omega_p^2)\nu}{A},$$

where $\omega_\infty = \sqrt{\omega_p^2 + \omega_{cr}^2 + \nu^2}$.

Figure 3b shows the minimum phase velocity versus ν/ω_{p1} at $H_0 = 1000$ Oe. The minimum phase velocity $v_{ph\,min}$ is seen to increase (i.e., $k_{x\,max}$ to decrease) with the collision frequency.

INHOMOGENEOUS MAGNETOPLASMA WAVES

Let us show that dispersion relation (2) has a solution on the complex plane $k_x = k_x' + ik_x''$. We assume that

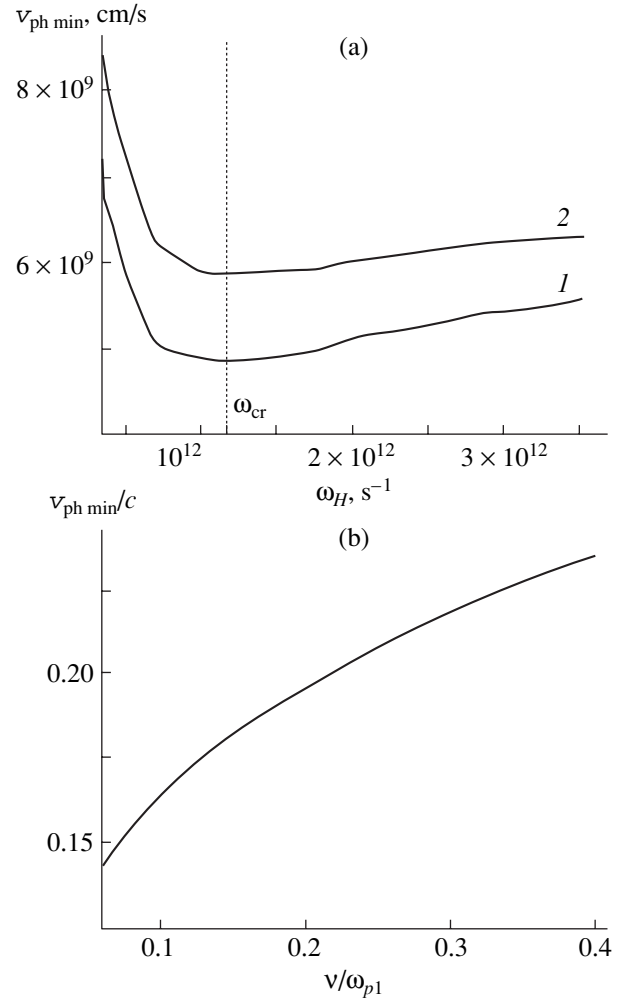$$k_x^2 \gg \frac{\omega^2}{c^2}|\varepsilon_{f1}, \varepsilon_2|.$$

In this case, dispersion relation (2) can be represented as

$$\cos \bar{k} d = \alpha + i\beta, \tag{7}$$

where

$$\alpha = \frac{1}{4\varepsilon_{\|1}\varepsilon_2}[((\varepsilon_{\|1} + \varepsilon_2)^2 + \varepsilon_\perp^2)\cos k_x'' d \cosh k_x' d$$

$$- ((\varepsilon_{\|1} - \varepsilon_2)^2 + \varepsilon_\perp^2)\cos k_x''(d_1 - d_2)\cosh k_x'(d_1 - d_2)],$$

$$\beta = \frac{1}{4\varepsilon_{\|1}\varepsilon_2}[((\varepsilon_{\|1} + \varepsilon_2)^2 + \varepsilon_\perp^2)\sin k_x'' d \sinh k_x' d$$

$$- ((\varepsilon_{\|1} - \varepsilon_2)^2 + \varepsilon_\perp^2)\sin k_x''(d_1 - d_2)\sinh k_x'(d_1 - d_2)]. \tag{8}$$

As follows from Eq. (7), $\cos \bar{k} d$ is real (which corresponds to the transmission band) if $\beta = 0$. The solution to this equation simultaneously satisfies two equalities:

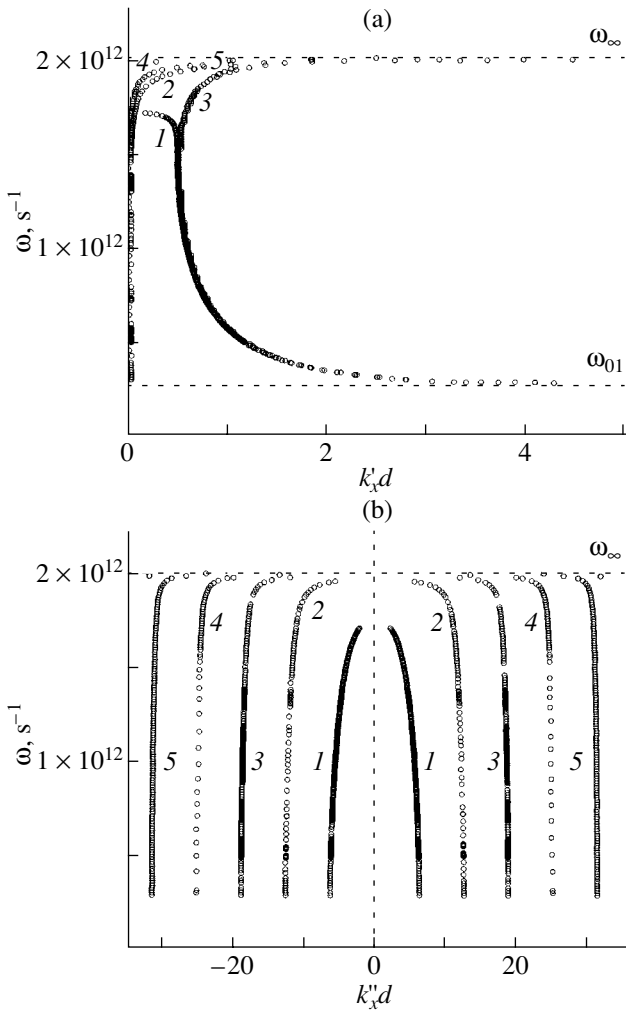$$k_x'' = \frac{\pi M}{d_1 + d_2} = \frac{\pi L}{d_1 - d_2}, \tag{9}$$

Fig. 4. Dispersion curves for inhomogeneous (complex) magnetoplasma waves.

Fig. 5. Dispersion curves for complex magnetoplasma waves with allowance for dissipation.

where $M, L = 0, \pm 1, \pm 2, \ldots$ and $M/L = d_1 + d_2/|d_1 - d_2|$ is a rational number.

The equation for k_x' can be shown to have a solution when M and L are both even or both odd. As follows from Eq. (9), the physical meaning of the existence condition for the existence of the complex modes is that the geometrical resonance be present in both layers simultaneously; i.e., the thickness of each layer must be a multiple of the half-wavelength [14]. Thus, the dispersion relation for magnetoplasma waves in an infinite semiconductor superlattice has many solutions on the complex plane k_x, which are determined by the numbers M and L. Figure 4 shows numerical solutions to dispersion relation (7) on the complex plane without allowance for collisions. The parameters of the periodic structure are the same as those in Fig. 1. Curves *1* refer to $M = 2$; curves *2*, to $M = 4$; curves *3*, to $M = 6$; and so on. It can be seen that a great number of complex modes occur in a periodic structure in a magnetic field. The
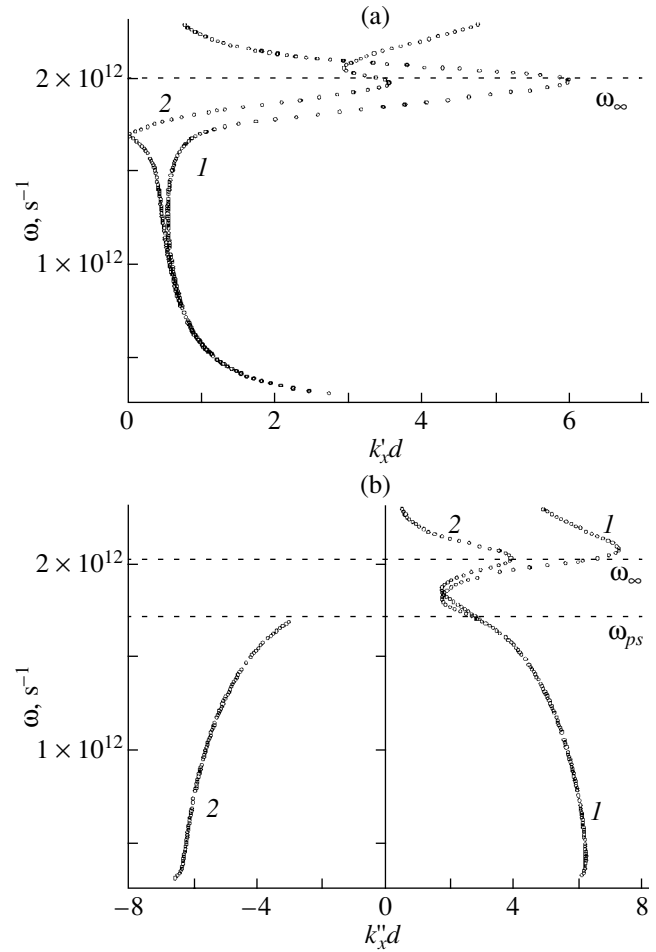
real parts of the wavenumbers ($\omega(k_x')$ curves in Fig. 4a) of different modes differ little from each other (for example, curves *2* and *4* and curves *1*, *3*, and *5* almost coincide). The $\omega(k_x'')$ dependence (Fig. 4b) has the form of a family of curves symmetric about $k_x'' = 0$; i.e., k_x'' takes two values $k_x'' = \pm |k_x''|$ for each mode. This is a distinctive feature of complex waves [12]. As the curves approach the asymptote $\omega = \omega_\infty$, $k_x' \longrightarrow \infty$ and $k_x'' \longrightarrow 0$.

Figure 5 shows dispersion curves for the complex magnetoplasma mode corresponding to $M = \pm 2$ at $\nu = 10^{11}$ s^{-1}. Let us compare this figure to Fig. 4. It can be seen that, when collisions are taken into account, dependences of the frequency on the real part of the wavenumber $k_x' d$ (Fig. 5a) for modes with positive (curve *1*) and negative (curve *2*) M do not coincide. The $\omega(k_x'')$ curves (Fig. 5b) are symmetric about $k_x'' = 0$ only

at frequencies below $\omega \approx \omega_{ps}$. At frequencies above $\omega \approx \omega_{ps}$, k_x'' becomes positive for both modes. Interestingly, the dispersion curve for positive M values turns into the curve for a normal magnetoplasma wave. These dispersion curves have a turning point at $\omega \approx \omega_\infty$, similar to those for usual magnetoplasmons.

CONCLUSIONS

The attenuation of electromagnetic waves in an infinite structure consisting of periodically alternating semiconductor and dielectric layers in a magnetic field has been studied theoretically. Dispersion characteristics are calculated with allowance for collisions and the finiteness of the speed of light in the semiconductor layers. Dissipation is shown to limit the phase velocity to a certain minimum value, which depends on the collision frequency.

The propagation of inhomogeneous (complex) waves and the effect of dissipation on their properties have also been studied. It is shown that, at high frequencies, the complex waves turn into normal waves.

The practical importance of this study is associated with the promise shown by structures with translation symmetry for developing passive and active microwave devices. In lattices with a period of a few micrometers to a few millimeters, the effects under study can be observed in the optical to centimeter wavelength range. In particular, the analysis of the dispersion properties of normal magnetoplasma waves propagating in such media shows that these waves can be applied in microwave filters and converters. These results can also be used in designing solid lasers.

ACKNOWLEDGMENTS

REFERENCES

1. A. C. Baynham and A. D. Boardman, Solid State Commun. **26**, 654 (1968).
2. A. C. Baynham and A. D. Boardman, J. Phys. C **2**, 619 (1969).
3. A. A. Bulgakov and Yu. E. Filippov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **28**, 1185 (1985).
4. R. F. Wallis, R. Szenics, J. J. Quinn, and G. F. Guiliani, Phys. Rev. B **36**, 1218 (1987).
5. R. F. Wallis and J. J. Quinn, Phys. Rev. B **38**, 4205 (1988).
6. M. S. Kushwaha, J. Phys. Chem. Solids **47**, 485 (1986).
7. A. A. Bulgakov and O. V. Shramkova, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 712 (2000) [Semiconductors **34**, 686 (2000)].
8. K. R. Alfano, J. Opt. Soc. Am. **60**, 66 (1970).
9. G. R. Kovner, K. W. Alexander, K. J. Beil, *et al.*, Phys. Rev. B **14**, 1458 (1976).
10. A. A. Bulgakov and Z. E. Eremenko, Opt. Spektrosk. **66**, 1094 (1989) [Opt. Spectrosc. **66**, 640 (1989)].
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
12. T. Tamir and A. Oliner, Proc. Inst. Electr. Eng. **110**, 311 (1963).
13. T. Tamir and A. Oliner, Proc. Inst. Electr. Eng. **110**, 325 (1963).
14. A. A. Bulgakov, S. A. Bulgakov, and M. Nieto-Vesperinas, Phys. Rev. B **58**, 4438 (1998).
15. F. G. Bass, A. A. Bulgakov, and A. P. Tetervov, *High-Frequency Properties of Semiconductors with Superlattices* (Nauka, Moscow, 1989).
16. A. I. Akhiezer, I. A. Akhiezer, and R. V. Polovin, *Plasma Electrodynamics* (Nauka, Moscow, 1974; Pergamon, Oxford, 1975).
17. A. A. Bulgakov and O. V. Shramkova, Radiotekh. Élektron. (Moscow) **46**, 236 (2001).

Translated by A. Khzmalyan

================= **RADIOPHYSICS** =================

Bulk and Surface–Bulk Magnetostatic Waves in Waveguides Produced by a Step Bias Field

A. Yu. Annenkov, S. V. Gerus, and S. I. Kovalev

Institute of Radio Engineering and Electronics (Fryazino Branch), Russian Academy of Sciences,
pl. Vvedenskogo 1, Fryazino, Moscow Oblast, 141190 Russia
e-mail: svg318@ms.ire.rssi.ru
Received July 14, 2003

Abstract—The propagation of magnetostatic waves in a ferromagnetic waveguide created by a step bias field is studied by numerical methods. Bias field configurations and frequencies are taken such that the width of the magnetic waveguide accommodates either only bulk waves or surface and bulk waves in combination. This work is an extension of earlier works, in which the propagation of surface waveguide modes was considered. © 2004 MAIK "Nauka/Interperiodica".

INTRODUCTION

Earlier investigation into magnetostatic waves (MSWs) propagating in magnetic waveguides (see, e.g., [1, 2]) did not reveal significant differences from wave propagation in a uniformly magnetized slab (the case considered by Damon and Eshbach [3]) as regards the dispersion curves. This paper and the previous one [4] consider bias field distributions that change noticeably the shape of MSW dispersion curves.

The bias field vector is aligned with the *z* axis and has a step discontinuity along this axis (Fig. 1). This field creates an in-plane magnetization of the ferromagnetic film *1*, which lies in the *y0z* plane. The field is uniform in the direction of the *y* axis, so that we are dealing with a magnetic waveguide *C* in this direction. This waveguide consists of two rectangular channels *A* and *B* with parameters h_A, H_A and h_B, H_B. MSW modes of the waveguide *C* can be considered as a hybridization of partial modes propagating in the rectangular channels *A* and *B*. Since their apertures are finite, the waveguides *A* and *B* are multimode.

Surface and bulk modes in each of the channels *A* and *B* are described by a family of dispersion curves. The frequency bands covered by these curves depend on the fields H_A and H_B in the channels, while the slopes of these curves are related to the channel widths h_A and h_B. These parameters were chosen so that the dispersion curves for bulk modes in channel *A* intersect with those for surface modes in channel *B*.

The bias field is assumed to be uniform along the *x* axis, which is an approximation. However, if a ferromagnetic film has a thickness *d* much smaller than the characteristic sizes h_A and h_B of the field nonuniformity

across the width of the waveguide (*z* axis), the nonuniformity along the thickness (*x* axis) can be neglected.

Our study of the properties of MSW modes in the waveguide is based on the solution of the Walker equation, which describes the distribution of the magnetostatic potential Ψ in terms of the dynamic permeability tensor $\mu(x, z)$:

$$\mathrm{div}[\mu(x, z)\mathrm{grad}\Psi] = 0. \qquad (1)$$

The nonuniformity of the field and that of the ferromagnetic film boundaries are included into the permeability tensor. We seek for a solution to the equation by the finite-difference method. A domain of interest is covered by a rectangular mesh. Then, Eq. (1) represented on this mesh in integral form is reduced to an equation written in terms of the quadratic matrix operator. The number of mesh cells is self-consistently obtained from the solution and its derivative. This method is described in more detail elsewhere [5, 6].
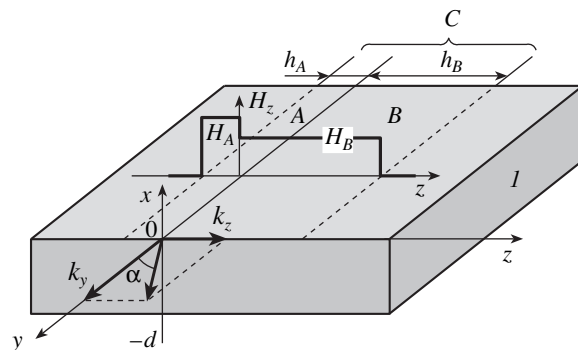

Fig. 1. Geometry of the problem: (*1*) ferromagnetic film and (*A, B, C*) magnetic waveguides.

The calculations result in dispersion curves for MSW modes propagating in the waveguide (Fig. 2) and distributions of the scalar potential $\Psi(x, z)$.

BULK MODES OF ONE CHANNEL

First, let us consider bulk MSWs in the rectangular channel A. The properties of bulk waves in an infinite in-plane magnetized ferromagnetic slab were discussed in [3]. It was shown that the distribution of the magnetic potential across the thickness (along the x axis) has the form of a standing wave that is described mathematically by a sine function whose argument (phase) varies from one plane to another. The distribution across the thickness is symmetric only when the MSW propagates parallel to the bias field (to the z axis); otherwise, it is asymmetric. With increasing wavenumber, the phase difference between the two surfaces for the first-order mode grows from 0 to π. Higher order modes differ from the first-order one by a multiple of half-waves fitted in the slab thickness. In the dispersion space $\omega(k_y, k_z)$, these modes form an infinite nest of trough-shaped surfaces. The lowermost surface refers to the first-order mode with the smallest phase difference across the film thickness, and the top-most surface tends to the plane $\omega_h = \gamma H_A$, where γ is the gyromagnetic ratio.

The behavior of bulk MSWs propagating in different directions relative to the bias field was studied in [7]. It was shown that there exists a direction in which the phase difference changes rapidly. The smaller the wavenumber, the faster this change: the phase changes roughly by π almost stepwise. Interestingly, this direction is defined by the angle $\alpha_c = \arctan\sqrt{4\pi M_0/H_A}$, which coincides with the cutoff angle of surface MSWs [3]. Here, M_0 is the saturation magnetization.

The existence of the magnetic channel along the y axis causes the wave to propagate only in the y direction. In addition, a standing wave forms across its width (in the z-axis direction). First, let us perform a simplified yet illustrative analysis of wave propagation in a channel with rigid magnetic walls, i.e., when the potential Ψ vanishes at the channel boundary. Such a channel guides a wave composed of two waves that are solutions to Eq. (1) for an infinite slab but have a fixed wavenumber k_z. These two waves produce an integer number of standing half-waves across the width of the channel. Dispersion curves for these modes are obtained by cutting the "troughs" described above by planes $k_z = k_{z0}n$, where n is an integer and k_{z0} corresponds to a single half-wave across the width of the channel. Modes with a different number of half-waves across both the channel width and the film thickness are produced. Although the initial dispersion troughs do not intersect, the dispersion curves obtained from different sections overlap, producing numerous intersections. As the number of half-waves accommodated on the channel width grows, the dispersion curves deform
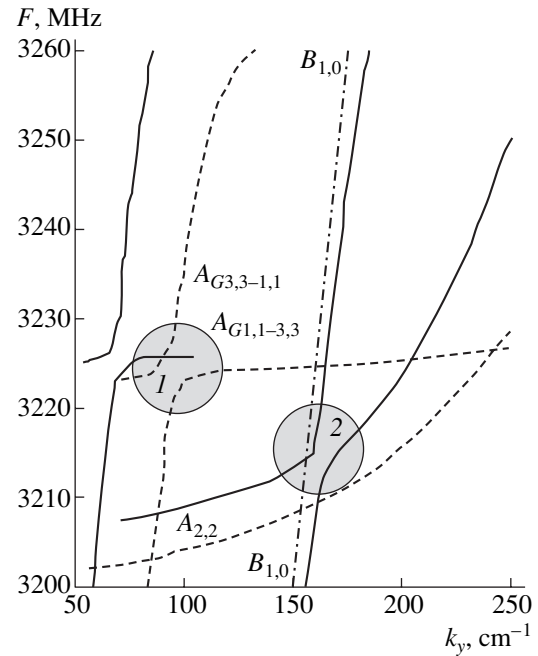

Fig. 2. Theoretical dispersion curves for the composite waveguide C and for its constituent channels A and B: solid lines, modes of the waveguide C; dashed lines, bulk modes $A_{m,n}$ and hybrid bulk modes $A_{Gm, n-p, q}$ of channel A; and dashed-and-dotted lines, surface modes $B_{m, 0}$ of channel B (m, n, p, and q are integers). Parameters of the magnetic waveguide: $h_A = 0.1$ mm, $h_B = 0.3$ mm, $H_A = 570$ Oe, $H_B = 445$ Oe, $4\pi M_0 = 1857.7$ G, and $d = 0.0183$ mm.

slightly, shifting towards greater wavenumbers, and the frequency band broadens. In contrast, an increase in the number of half-waves across the film thickness shifts the dispersion curves in the opposite direction.

Let us mark bulk waves by two subscripts: one for the number of half-waves across the channel width and the other for the number of half-waves across the film thickness (for example, $A_{3, 2}$ means that this mode propagates in channel A and has three half-waves across the channel width and two half-waves across the film thickness).

When the wavenumber k_y is varied from zero to infinity, the direction of the total wavenumber $\mathbf{k}_y + \mathbf{k}_z$, which refers to an MSW in an infinite slab and is initially oriented transversely to the channel, becomes parallel to the channel. As follows from [7], the phase difference is bound to increase with thickness slowly except for the k_y range where the wave goes through the critical angle α_c. In this range, the phase difference rapidly grows from a minimum to a maximum value. The smaller $|\mathbf{k}_y + \mathbf{k}_z|$ near the critical angle, the faster the variation of the phase difference with thickness.

Now, let us consider the behavior of bulk waves in a nonideally rigid channel. To this end, we seek for a numerical solution to Eq. (1) in a region whose boundaries are far away from the boundaries of the waveguide, so that the potential Ψ at the boundaries of the region vanishes. As a result, the potential at the boundaries of channels A and B remains finite. Because of this, the wave field outside the channel decreases and the wave distribution across the width of the channel becomes nonsinusoidal. The initial modes are no longer orthogonal and may interact at the point where the dispersion curves intersect. Because of this, the marking of the modes by the number of half-waves accommodated on the channel width and on the film thickness, which was used in the case of rigid walls, becomes incorrect. A particular mode may contain a different number of half-waves over both the channel width and the film thickness, depending on the frequency.

Our calculations show that various modes behave variously at the points of intersection: some of them repel each other, and others do not. The absence of repulsion means that the modes do not interact. The modes do not interact if the integral of the product of the distributions of their magnetostatic potentials (hereafter simply potentials) equals zero. Let us denote the potentials of two modes as $\Psi_1(x, z)$ and $\Psi_2(x, z)$. Then, their interaction integral can be written as

$$\iint \Psi_1(x, z)\Psi_2(x, z)\,dx\,dz.$$

The integral vanishes if the integrand is antisymmetric about at least one coordinate. Since the thickness distribution of the potential is, in general, asymmetric, we will consider the distribution across the width (along the z axis). It is symmetric (antisymmetric) if the number of half-waves is odd (even). The product of a purely antisymmetric mode by a symmetric mode yields an antisymmetric integrand. Thus, at points where even and odd modes intersect, mode interaction and, hence, the repulsion of the dispersion curves are not expected.

Figure 2 shows the dispersion curves for bulk waves in channel A (dashed lines). The dispersion curves for the modes $A_{1,1}$ and $A_{3,3}$ (omitted in the figure) repel each other at the point of interaction (circle *1*) and produce two hybrid modes, $A_{G1,1-3,3}$ and $A_{G3,3-1,1}$. Note that here the subscripts indicate how many and which of the original modes from the rigid-wall channels produce hybrid waves in the soft-wall channel. At the same time, the dispersion curve for the mode $A_{2,2}$ intersects the dispersion curve for the mode $A_{G1,1-3,3}$ without repulsion. This corroborates our conclusion that interaction does not occur between modes that contain even and odd numbers of half-waves over the channel width.

In the region of repulsion of the dispersion curves, the distribution of the potential gradually changes. Figure 3 demonstrates the transformation of the magnetostatic potential in the region where the modes $A_{1,1}$ and $A_{3,3}$ repel each other. For surface waves [4], the number of half-waves accommodated across the width of channel C for a particular mode is constant and equals its index, while, for the bulk hybrid modes $A_{G1,1-3,3}$ and $A_{G3,3-1,1}$, the number of half-waves varies along their dispersion curves. On different portions of the curves, either one half-wave or three half-waves are accommodated both across the width of the channel and across the film thickness.

Consider the behavior of the phase difference on the example of the mode $B_{2,2}$. Figure 4a shows the phase difference across the channel width and film thickness versus the wavenumber k_y. To calculate the phase difference, the actual potential distribution was approximated by a sine function. The phase difference across the film thickness (curve *1*) is seen to grow rapidly from a minimum to a maximum in the interval $k_y = 200$–350 cm^{-1}.

The distribution of the potential across the thickness is asymmetric [7]. On the lower surface, the phase varies slowly with the wavenumber and is close to zero, which implies weak fields at the boundary and a weak leakage of the wave from the film through this surface. On the upper surface, the phase is roughly equal to the phase difference and changes sharply near the critical angle α_c. Both before and after the critical angle, the phase is close to a multiple of π. As in the case of the lower surface, this means that the fields at the boundary and outside the slab are low. However, when the angle takes the critical value, the field at the boundary reaches a maximum at a phase of $n\pi + \pi/2$. At this point, the leakage of the field from the slab is the highest (curve b_0 in Fig. 4b). It is of interest that the phase difference versus the integer number of half-waves across the channel width is the smallest at this point (Fig. 4a, curve *2*). The potential at the channel boundaries, as well as the field, reaches a maximum at this point, which means a maximal leakage of the field from the channel into the slab (curve b_0 in Fig. 4c). Thus, the dispersion curve for the bulk wave has a point where the field leakage out of the channel is the highest in both coordinates and which corresponds to the cutoff angle of the total wave vector $\mathbf{k}_\tau = \mathbf{k}_y + \mathbf{k}_z$.

MODES OF THE COMPOSITE WAVEGUIDE

Now consider modes that propagate in the composite waveguide C. They result from interaction between bulk modes in the narrow channel A and surface modes in the wide channel B. In Fig. 2, the modes of the composite channel are indicated by the solid lines; bulk modes of the narrow channel, by the dashed lines; and
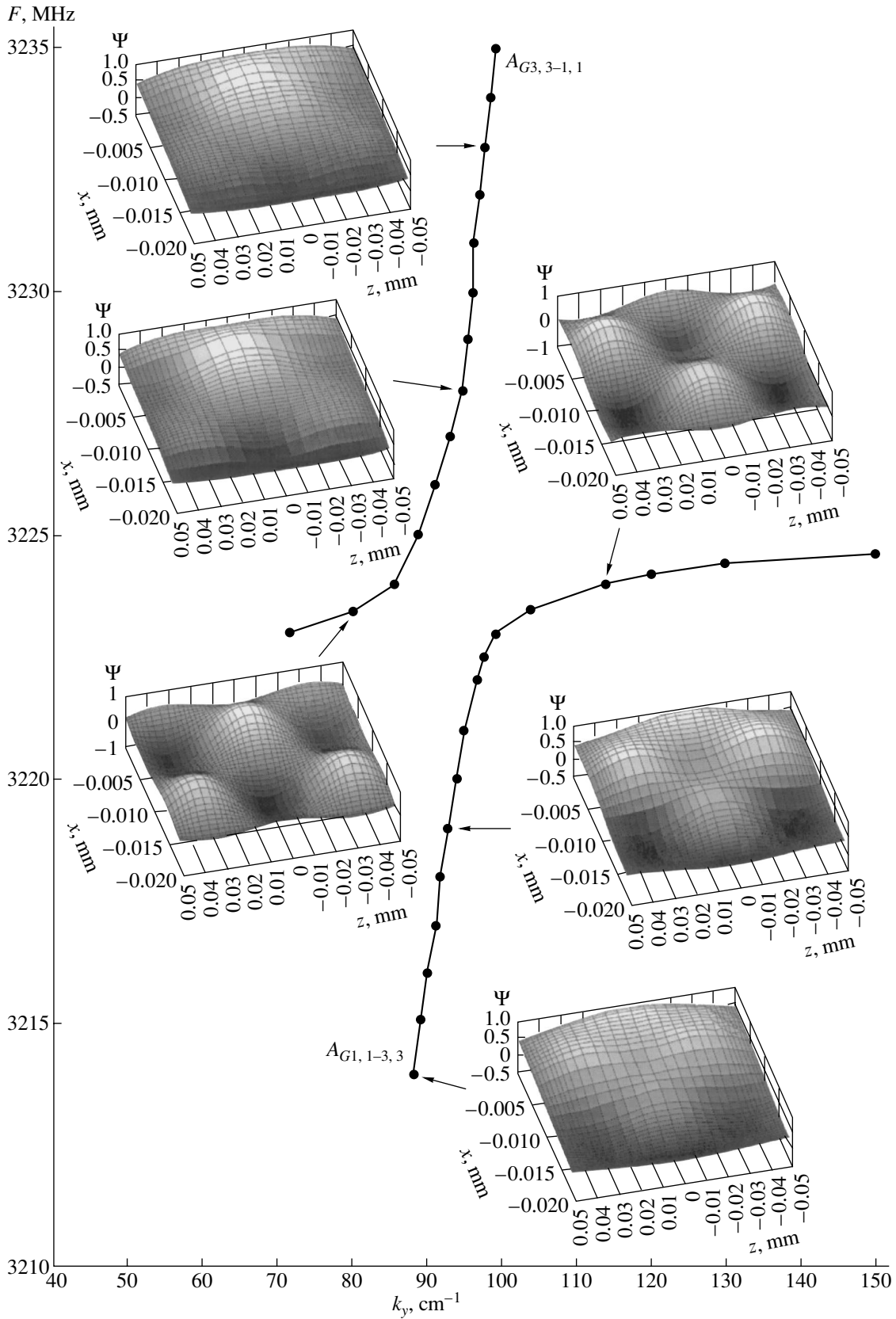
Fig. 3. Distribution of the potential Ψ in the region of repulsion of the bulk waves $A_{1,1}$ and $A_{3,3}$ (the region is indicated by circle *1* in Fig. 2).
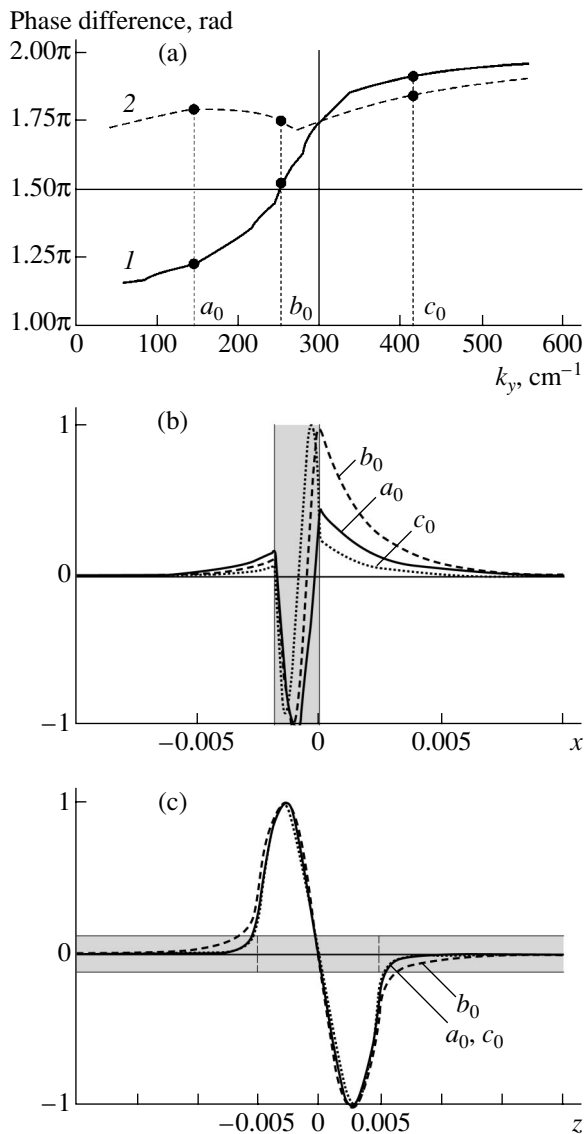
Phase difference, rad

Fig. 4. (a) Phase difference across (1) the film thickness and (2) the channel width vs. wavenumber for the bulk mode $A_{2,2}$ and the distribution of the magnetic potential Ψ across (b) the thickness and (c) the width for the points a_0 ($\alpha < \alpha_c$), b_0 ($\alpha = \alpha_c$), and c_0 ($\alpha > \alpha_c$).

surface modes of the wide channel, by the dashed-and-dotted lines.

For the composite waveguide, the regions of repulsion of the dispersion curves can be subdivided into two types. The first type embraces regions where the dispersion curves for modes in different channels intersect. This case is similar to the one discussed in [4], where the tails of the wave functions leave the channels A and B and interact, causing the modes of both channels to join together at the common boundary. The second type covers regions of repulsion, where the dispersion curves for bulk modes of the narrow channel A intersect (see the previous section). In this case, interaction

between different modes of the narrow channel results in the redistribution of the magnetostatic potential in the channel.

As follows from the plots, interaction between the channels displaces significantly the dispersion curves of the composite waveguide relative to the curves for each of the constituent channels. The ranges of modes of the narrow channel shift toward smaller wavenumbers (higher frequencies), whereas the ranges that refer to surface modes of the wide channel move toward higher wavenumbers.

Interaction between modes of the narrow channel A was considered above. The presence of the second channel changes the overall pattern insignificantly. Therefore, we will focus on the interaction between modes in different channels.

Figure 5a shows the dispersion curves that appear when the bulk mode $A_{2,2}$ of the narrow channel and the surface mode $B_{1,0}$ of the wide channel repel each other, and Fig. 5d demonstrates the three-dimensional distributions of the magnetostatic potential for six points of these curves. For modes of the composite waveguide, the magnetostatic potential amplitude is greater in the channel near whose dispersion curve the dispersion plot of the composite waveguide passes at a given point. At the point of repulsion, the wave amplitude is gradually transferred from one channel to the other, which is clearly seen in Fig. 5d. The potential distribution for the points a_1, b_1, and c_1, which belong to the left-hand dispersion curve, shows that the amplitude of the bulk mode $A_{2,2}$ of the narrow waveguide decreases and the amplitude of the mode $B_{1,0}$ of the wide channel increases. For the dispersion curve on the right (the points a_2, b_2, and c_2), the reverse is true: the amplitude of the bulk mode $A_{2,2}$ of the narrow waveguide increases, while the amplitude of the mode $B_{1,0}$ of the wide channel decreases.

It was noted [7] that surface modes in both channels may join together at the boundary in two ways, each corresponding to one of the curves produced at the point of repulsion. One way is in-phase matching, where half-waves at the boundary are of the same polarity and join together without intersecting the zero axis. The other is referred to as antiphase matching, where half-waves of different polarity join together. In the latter case, we have one more intersection with the zero axis than in the case of in-phase matching. For surface modes, the mode index remains unchanged throughout the dispersion curve. This approach fails in joining surface and bulk modes, since the phase of a bulk wave at the channel boundary depends on the thickness coordinate; accordingly, for the same mode, joining may be either in-phase or antiphase, depending on x (see Figs. 5b, 5c). Nevertheless, if joining is
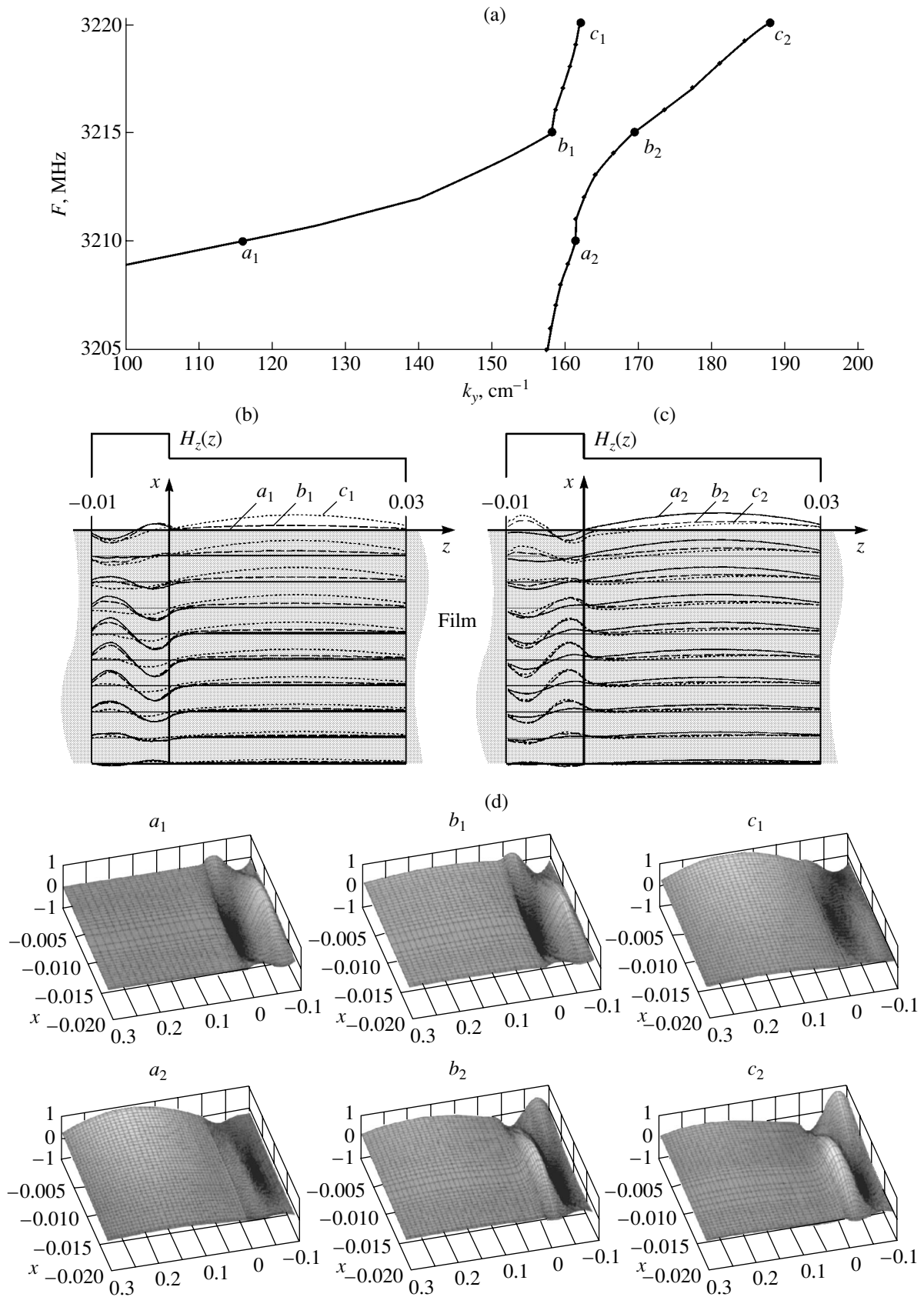
Fig. 5. Distribution of the potential Ψ in the region where the bulk mode $A_{2,2}$ of the narrow channel and the surface mode $B_{1,0}$ of the wide channel repel each other (the region is indicated by circle 2 in Fig. 2).

in-phase for the left-hand dispersion curve, it is antiphase for the right-hand curve (and vice versa) if we consider the same section across the width.

CONCLUSIONS

It is shown that the dispersion curves of bulk modes in a rectangular channel intersect. Some of these modes interact, creating regions of repulsion; others do not interact and intersect each other without repulsion.

At the sites of repulsion of bulk waves, the number of magnetostatic potential half-waves varies along the dispersion curve.

The dispersion curve for a bulk mode has a point at which the field leakage through the channel boundary and film surface is maximum. This point corresponds to the characteristic cutoff angle of the total wave vector, which coincides with the cutoff angle of the surface MSWs.

Regions where modes of the composite waveguide repel each other are those where dispersion curves of either modes in different channels or bulk modes in one channel intersect.

REFERENCES

1. M. Miyazaki, K. Yashiro, and S. Ohkawa, IEEE Trans. Microwave Theory Tech. **33**, 421 (1985).

2. A. Yu. Annenkov, S. V. Gerus, and S. I. Kovalev, Zh. Tekh. Fiz. **68** (2), 91 (1998) [Tech. Phys. **43**, 216 (1998)].

3. R. W. Damon and J. R. Eshbach, J. Phys. Chem. Solids **19**, 308 (1961).

4. A. Yu. Annenkov, S. V. Gerus, and S. I. Kovalev, Zh. Tekh. Fiz. **72** (6), 85 (2002) [Tech. Phys. **47**, 737 (2002)].

5. I. V. Vasil'ev and S. I. Kovalev, Radiotekh. Élektron. (Moscow) **38**, 2174 (1993).

6. A. Yu. Annenkov, I. V. Vasil'ev, S. V. Gerus, and S. I. Kovalev, Zh. Tekh. Fiz. **65** (4), 71 (1995) [Tech. Phys. **40**, 330 (1995)].

7. A. Yu. Annenkov and S. V. Gerus, Zh. Tekh. Fiz. **69** (1), 82 (1999) [Tech. Phys. **44**, 74 (1999)].

Translated by A. Khzmalyan

=========== SURFACES, ELECTRON ===========
AND ION EMISSION

Thermal Destruction of Two-Dimensional Graphite Islands on Refractory Metals (Ir, Re, Ni, and Pt)

N. R. Gall', N. P. Lavrovskaya, E. V. Rut'kov, and A. Ya. Tontegode

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: gall@ms.ioffe.ru
Received July 10, 2003

Abstract—Thermal destruction of two-dimensional graphite films on Ni(111), Re(10–10), Ir(111), and Pt(111) substrates is studied. It is shown that the detachment of an edge carbon atom from an island is a limiting process stage for all the cases. The activation energy of this process varies from 2.5 eV for nickel to 4.5 eV for iridium. The variation of the activation energy is associated with the ability of the metal surface to form strong chemisorptive bonds with valence-active edges of graphite islands, which loosen C–C bonds in graphite. © 2004 MAIK "Nauka/Interperiodica".

Two-dimensional graphite films (TGFs) on solids, which can be referred to two-dimensional crystals, are an interesting scientific object. These films form on many metals (Ir, Pt, Rh, Ni, Re, Mo, and Ru) without crystallographic match between the graphite layer and the metal surface, excluding the (111)Ni face [1, 2]. The graphite network is clearly observable when a graphite monolayer on (111)Ir or (10–10)Re is examined by scanning tunnel microscopy and atomic force microscopy [3, 4]. Similar results were obtained by low-energy electron diffraction (LEED) [5, 6], SEX-AFS [7], and inelastic electron scattering [8, 9]. Having saturated valence, a two-dimensional graphite layer is bonded to a metal surface by weak van der Waals forces [1]. Because of this, TGFs retain all physical and chemical properties inherent to the basal plane of a graphite single crystal, such as the work function, chemical and catalytic passivity, adsorptivity, etc. [2].

Earlier [2], we studied in detail the formation, growth, and physicochemical properties of graphite films on various refractory metals. In this paper, we consider other issues responsible for the physicochemical properties of TGFs, namely, the mechanisms, kinetics, and energy of thermal destruction of two-dimensional graphite islands on (111)Pt, (10–10)Re, (111)Ni, and (111)Ir surfaces.

EXPERIMENTAL

The mass spectrometric technique applied in this work was described elsewhere [10]. We also used an ultra-high-vacuum high-resolution prism Auger electron spectrometer [11], which allows one to identify the chemical state of surface carbon (adsorbed carbon "gas," surface carbide, graphite, diamond, or fullerite) by the form and energy position of the C_{KVV} Auger spectrum [12]. The Auger spectra from graphite islands

and the basal plane of a graphite single crystal are similar. Samples used were thin textured metal strips measuring $40 \times 1 \times 0.02$ mm. The strips were heated directly by passage of alternating current and decontaminated by high-temperature heating in oxygen combined with ultra-high-vacuum annealing by the technique described in [2]. The texture was observed on the (111) face for Ni, Ir, and Pt strips and on the (10–10) face for Re strips. The surfaces of the strips were uniform in work function value. The strip temperature was measured by a micropyrometer or by linearly extrapolating the dependence of the sample temperature on the filament current (in the range where the pyrometer fails). The spread of sample temperatures measured pyrometrically was within $\Delta T = 5$ K.

Graphite film can be applied on metals by several methods. One is flux deposition of carbon atoms on a heated metal surface using a special absolutely calibrated source [13]. Another method is cracking of hydrocarbon molecules (usually benzene C_6H_6) on a heated metal. In this case, the molecule dissociates into atoms, hydrogen is desorbed, and carbon alone remains on the surface. Both methods lead to the growth of graphite islands on the surface and require knowledge of the temperature range ΔT of island formation [2]. Merging together, the graphite islands form a solid graphite film.

These islands are two-dimensional (see reviews [1, 2]). Note that the use of benzene provides a uniform graphite monolayer automatically, since benzene molecules do not dissociate on the passive graphite surface and only a monolayer film can be formed [1]. The TGF growth, continuity of the films, and their thickness were studied with high-resolution Auger spectroscopy [11].

In this paper, the relative area $\Delta S/S$ of graphite islands was determined by a simple and still reliable

method, namely, by probing the surface by a low-density flux of CsCl molecules (for details, see [1]). This method relies on the fact that CsCl dissociates on valence-passive graphite islands only insignificantly. At the same time, at $T > 800$ K, the dissociation CsCl $\longrightarrow$ Cs + Cl proceeds with a probability of 100% on the parts of the surface that are free of graphite islands (pure metal, areas with chemisorbed carbon gas, or metal carbide). For $T \geq 800$ K, cesium, one of the reaction products, is desorbed from the surface in the form of Cs^+ ions produced by surface ionization [14]. For all the metals and film systems used in this work, the ionization potential of cesium ($V_{Cs} = 3.89$ eV) was invariably lower than the work function of graphite-uncovered areas of the surface; the cesium is completely desorbed in the form of Cs^+ ions [14]. Therefore, measuring the cesium ion current $I^+ = e\nu(S_{tot} - S)$, we found the fraction of the surface that is free of graphite islands (here, ν is the flux density of CsCl molecules reaching the surface, S is the surface area occupied by the islands, and S_{tot} is the total area of the sample). If the cesium ion current from the initially pure surface is measured to be I_0^+ ($I_0^+ = e\nu S_{tot}$), we have

$$\frac{I^+}{I_0^+} = \frac{S_{tot} - S}{S_{tot}} = 1 - \frac{S}{S_{tot}} \quad \text{or} \quad S/S_{tot} = 1 - \frac{I^+}{I_0^+}.$$

In other words, to find the occupied fraction of the surface, one must merely measure two currents of cesium ions: I_0^+ and $I^+ = f(S)$. Obviously, the current I^+ measured at any time will bear information on the relative surface area occupied by graphite islands. In experiments, we used low-density fluxes of CsCl molecules, $\nu_{CsCl} = 1 \times 10^{10}$ cm^{-2} s^{-1}, and could quickly cut off the flux toward the sample surface. Thus, the dependence $I^+ = f(t)$ reflects the kinetics of growth or destruction of the graphite islands.

RESULTS

The formation of graphite islands on a metal surface at a temperature T implies that there is a higher temperature $T_d > T$ at which the islands start to be destroyed. The experiments showed that, for all the four metals, studied the temperature T_d at which graphite islands are destroyed is much lower than the temperature of mechanical destruction of the substrate and the temperature T_{des} at which noticeable desorption of carbon atoms from a refractory metal surface is observed (usually $T_{des} \geq 2000$ K). The destruction of graphite islands on platinum, nickel, and rhenium decreases the total amount of carbon on the surface. This is apparently because carbon atoms detached from the island edge dissolve in the metal. The only exception is iridium, where carbon hardly dissolves at $T \leq 2000$ K and remains on the surface as chemisorbed carbon gas [10] after the destruction of the islands.

We assume that the elementary act of destruction is the detachment of a carbon atom from the island edge and its transition into the chemisorbed state on the metal surface. Suppose, for simplicity, that graphite islands are disks of equal surface area S_1. Then, as the temperature increases to $T > T_d$, the number n_1 of carbon atoms in a single island decreases with time by the law

$$-\frac{dn_1(t)}{dt} = n_b(t)W = n_b(t)C\exp[-E_{sep}/kT]. \quad (1)$$

Here, $n_b(t)$ is the number of carbon atoms in the boundary layer of the island at a time t, W is the probability of two-dimensional sublimation, C is the preexponential, and E_{sep} is the activation energy of two-dimensional sublimation (in other words, the energy of separation of an edge carbon atom from the island).

For simplicity, we also suppose that carbon atom packing in the islands is identical to that in the graphite basal plane and is independent of the island size; then, $S_1 \approx an_1$, where a is the surface area occupied by a carbon atom. Assuming that $n_b = 2\pi r/d$ (where r is the radius of a carbon island (disk) and $d = \sqrt{4a/\pi}$ is the effective diameter of a carbon atom in the graphite network) and taking into account that

$$r = \sqrt{\frac{n_1 a}{\pi}} = \sqrt{\frac{aN}{\pi m}}$$

(where N is the total number of carbon atoms in all islands per 1 cm^2 and m is the island density), we obtain

$$n_b = \frac{\pi}{\sqrt{m}}\sqrt{N(t)}.$$

From Eq. (1), we then find

$$\frac{dN}{\sqrt{N}} = -\pi\sqrt{m}C\exp[-E_{sep}/kT]dt. \quad (2)$$

Under the initial condition $N = N_0$ at $t = 0$, the solution to Eq. (2) for arbitrary t has the form

$$1 - \sqrt{\frac{N}{N_0}} = -\frac{\pi\sqrt{m}C\exp\left[-\dfrac{E_{sep}}{kT}\right]}{2\sqrt{N_0}}t. \quad (3)$$

The ratio $N(t)/N_0 = S(t)/S_0$, where S_0 is the area occupied by the islands at $t = 0$, is determined experimentally by the method of CsCl molecule dissociation mentioned above. Having a set of curves (1 − $\sqrt{S(t)/S_0}$) = $f(t)$ at hand, we can find the energy of carbon atom detachment from the graphite island using the dependence $\ln t = f(1/kT)$ for every S = const. Knowing E_{sep}, one can also determine the value of $\sqrt{m}C$ from Eq. (3).

As an example, let us consider the destruction of graphite islands on the (10–10) face of rhenium. A pure rhenium sample was kept for a short time (t = 10–30 s) in benzene vapors ($P_{C_6H_6} \sim 1 \times 10^{-7}$ Torr) at T = 1050 K. In this case, the surface was rapidly covered by graphite islands, while the interior of the metal remained almost intact. It turned out that carbon at this temperature is present on the surface only in the form of graphite islands, because individual carbon atoms dissolve rapidly in the metal [15]. Nevertheless, for the experiment to be valid, we annealed the sample for several minutes at T = 1050 K after benzene had been evacuated. The graphite islands were not destroyed in this case, and a small (if any) amount of carbon accumulated in the near-surface layer was distributed over the metal volume. Proceeding in this way, we decreased the probability of the reverse (from the volume to the surface) carbon flux. Next, the rhenium temperature was increased to $T_d \geq 1200$ K and the kinetics of island destruction was recorded; i.e., the dependence $S/S_0 = f(t)$ was constructed. In this case, the islands were completely destroyed (Fig. 1). Note that the destruction curves were reproduced well from test to test though the sample volume was not decontaminated from the carbon accumulated. This suggests that, at these temperatures, the carbon concentration in the bulk of rhenium is far from the limiting solubility, so that the carbon diffuses rapidly over the metal lattice and the reverse carbon flux can be neglected.

The data obtained were processed as follows. Using the curves in Fig. 1, we plotted the dependences $\ln t = f(1/kT)$ for different S = const to determine the energy of carbon atom separation from an island $E_{sep} = (3.0 \pm 0.2)$ eV. In the experiments, S/S_{tot} was varied in the range 0.2–0.6, while $S/S_{tot} = 1$ corresponded to a graphite monolayer with $N_c = 3.86 \times 10^{15}$ cm^{-2}. Then, we constructed the dependence

$$S(t)/S_0 = (1 - Bt)^2,$$

where

$$B = \frac{\pi\sqrt{m}\,C\exp[-E_{sep}/kT]}{2\sqrt{N_0}}. \qquad (4)$$

The value of B was chosen such that the calculated dependences $S/S_0 = f(t)$ agreed with the experimental ones (Fig. 1, dashed lines). Knowing B and using Eq. (4), one can determine the product of the preexponential by the square root of the number of the islands, $\sqrt{m}\,C = 3 \times 10^{18}$ cm^{-1} s^{-1}. In our previous studies of the graphite monolayer topography on (10–10)Re by scanning tunneling microscopy, we estimated the density of graphite islands in near-monolayer coatings. It was found to be $m \sim 10^{12}$ cm^{-2} [3, 4]. Knowing m, one finds that $C \sim 10^{12}$ s^{-1}.
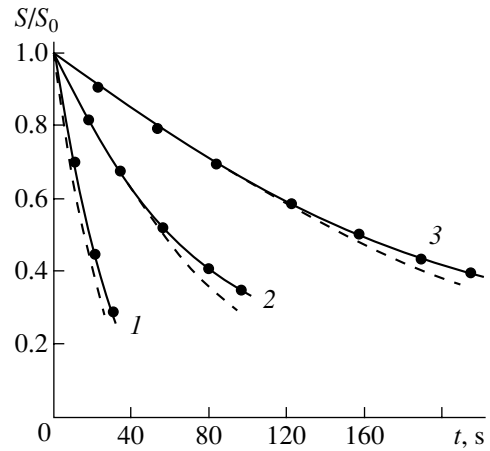
Fig. 1. Decrease in the relative area occupied by graphite islands on (10–10)Re upon isothermal annealing at a temperature of (*1*) 1280, (*2*) 1240, and (*3*) 1215 K. The initial ratio is $S/S_{tot} = 0.3$ for all the cases.

Destruction of graphite islands on (111)Ni and (111)Pt was studied in a similar way. The difference was solely in the temperatures of island destruction (see table). The values of $\sqrt{m}\,C$ were of the same order of magnitude as in the case of rhenium.

On iridium, graphite islands start to be destroyed at $T_d \geq 1600$ K. This means that the binding energy of an edge carbon atom with an island is much higher than on the substrates considered above. Since carbon atoms detached from the island do not dissolve in the iridium and are not desorbed, the total carbon concentration on the surface remains constant and the concentration of carbon atoms in the chemisorbed carbon gas increases when the surface area of the graphite islands decreases upon their thermal destruction. This, in turn, enhances the flux of carbon atoms from the chemisorbed gas to the graphite islands. Therefore, the final result of destruction for a given temperature T_d = const is dynamic equilibrium between the flux of two-dimensional sublimation and the carbon flux to the islands from the chemisorbed gas. In this case, the island surface area stabilizes at some temperature-dependent level (Fig. 2). If the temperature drops to $T < 1600$ K, the island area takes the previous value $S = S_0$; i.e., all the carbon gas passes again into the graphite islands. To

Energies of activation of edge carbon atom detachment from the graphite island, E_{sep}, and temperature ranges ΔT_d of graphite island destruction on different substrates

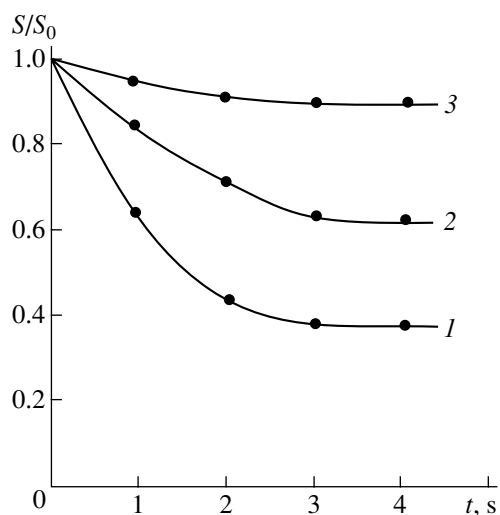| Me	E_{sep}, eV	ΔT_d, K
(111)Ni	2.5	950–1050
(10–10)Re	3.0	1200–1300
(111)Pt	3.2	1300–1400
(111)Ir	4.5	1650–1850

Fig. 2. Decrease in the relative area occupied by graphite islands on (111)Ir upon isothermal annealing at a temperature of (*1*) 1940, (*2*) 1840, and (*3*) 1715 K. The initial ratio is $S/S_{tot} = 0.5$ for all the cases.
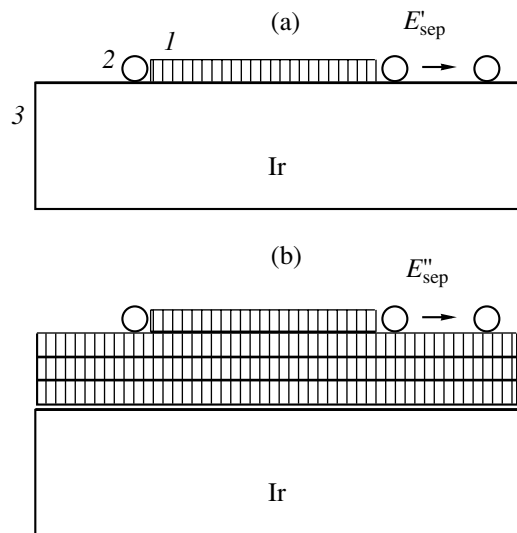

Fig. 3. Schematic of thermal destruction of two-dimensional graphite islands. (a) Iridium surface; E'_{sep} is the activation energy of detachment of an edge carbon atom contacting with the metal. (b) Multilayer graphite film or graphite single crystal; E''_{sep} is the activation energy of edge atom detachment from the graphite layer. (*1*) Graphite island, (*2*) edge carbon atom in the graphite network, and (*3*) iridium substrate.

exclude the effect of the reverse carbon flux from the chemisorbed gas, only the earliest portions of the dependences $S/S_0 = f(t)$ were processed. The experimental results are also presented in the table.

DISCUSSION

It is seen from the table that, in the case of graphite islands adsorbed on metal surfaces, the separation energy of an edge carbon atom in the layer may be much lower than that for bulk graphite (~6 eV at $T_d >$ 2300 K) [1]. It is reasonable to assume that valence-active edges of graphite islands produce strong chemisorbed bonds with atoms of the metallic substrate. This loosens C–C bonds between an edge carbon atom and its neighbors and, accordingly, decreases markedly the temperature of graphite island destruction on the metal (table; Fig. 3a). As follows from the experiment, graphite islands on a nickel are the easiest to destroy. This correlates well with the fact that nickel is the only carbide-forming metal of all the metals studied. The islands on iridium offer the highest thermal stability. Iridium does not form carbides, dissolves carbon in negligible amounts, and is hardly capable of forming C–Me chemical bonds. Interestingly, if a graphite film with a thickness of several atomic layers is formed on (111)Ir by deposition in an atomic flux at $T \approx 1700$ K [2] (Fig. 3b), the temperature of graphite layer destruction (and, accordingly, of carbon removal from the surface) rises to $T_d > 2300$ K. This temperature corresponds to $E_{sep} \approx 6$ eV, i.e., to the value typical of thermal destruction of bulk graphite. In fact, in this case, there is no contact between an edge atom of the top graphite layer and the metal; hence, there are no reasons to expect that the activation energy of edge atom detachment from a graphite island will decrease because of chemisorbed C–Me bonds.

CONCLUSIONS

In this paper, a destruction mechanism for graphite islands on a metal is suggested and the binding energy of edge carbon atoms with a graphite island is determined. This energy is responsible for the thermal destruction of islands on Ni(111), Re(10–10), Ir(111), and Pt(111) substrates. Graphite islands on iridium, which does not form carbides, offer the highest temperature stability: the binding energy of an edge carbon atom with an island reaches 4.5 eV ($T_d > 1600$ K). Unique potentialities of an original method for analyzing a surface containing graphite islands, the method of CsCl molecule dissociation, are demonstrated. Unlike many other traditional methods of surface diagnostics, this simple method is sensitive only to the graphite phase of surface carbon and allows one to keep track of the dynamics of growth and destruction of graphite islands at any temperatures above 850 K. It seems that this method may also be used to study other films with saturated valence.

REFERENCES

1. A. Ya. Tontegode, Prog. Surf. Sci. **38**, 201 (1991).

2. N. R. Gall, E. V. Rut'kov, and A. Ya. Tontegode, Int. J. Mod. Phys. B **11**, 1865 (1997).

3. I. V. Makarenko, A. N. Titkov, V. Vakar, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **40**, 1570 (1998) [Phys. Solid State **40**, 1423 (1998)].

4. I. V. Makarenko, A. N. Titkov, V. Vakar, *et al.*, Poverkhnost, No. 7, 39 (1999).

5. M. Eizenberg and J. M. Blakely, Surf. Sci. **82**, 228 (1979).

6. Zi-Pu Hu, D. F. Ogletree, M. A. van Hove, and G. A. Somorjai, Surf. Sci. **180**, 433 (1987).

7. R. Rossei, M. de Crescenzi, F. Sette, *et al.*, Phys. Rev. B **28**, 1161 (1983).

8. T. Aizawa, R. R. Souda, S. Otani, *et al.*, Phys. Rev. B **42**, 11469 (1990).

9. T. Aizawa, Y. Hwang, W. Hayami, *et al.*, Surf. Sci. **260**, 311 (1990).

10. N. A. Kholin, E. V. Rut'kov, and A. Ya. Tontegode, Surf. Sci. **139**, 155 (1984).

11. V. N. Ageev, E. V. Rut'kov, A. Ya. Tontegode, and N. A. Kholin, Fiz. Tverd. Tela (Leningrad) **23**, 2248 (1981) [Sov. Phys. Solid State **23**, 1315 (1981)].

12. E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Phys. Rev. Lett. **74**, 758 (1995).

13. N. R. Gall, E. V. Rut'kov, A. Ya. Tontegode, *et al.*, J. Chem. Vap. Deposition **6** (1), 62 (1997).

14. E. Ya. Zandberg and N. I. Ionov, *Surface Ionization: Israel Program for Scientific Translations* (Jerusalem, 1971).

15. N. R. Gall, S. N. Mikhailov, E. V. Rut'kov, and A. Ya. Tontegode, Surf. Sci. **191**, 185 (1987).

Translated by M. Astrov

SURFACES, ELECTRON
AND ION EMISSION

Formation of Rings around a Field-Emission Image and Possible Applications of This Effect

K. N. Nikolski, A. S. Baturin, A. I. Knyazev, R. G. Tchesov, and E. P. Sheshin

Moscow Institute of Physics and Technology, Institutskiĭ pr. 9, Dolgoprudnyĭ, Moscow Oblast, 141700 Russia
e-mail: nkn@lafeet.mipt.ru
Received March 25, 2003; in final form, July 30, 2003

Abstract—By using the method of pulsed field electron emission, it is shown that, as the current density approaches the ultimate value, there appears a bright ring around a normal field-emission image. This effect is frequently observed in the case of planar field-emission cathodes. It is suggested that secondary electrons emitted from the anode return to the anode under the action of the electric field, producing the rings. Field-emission applications based on this effect are discussed. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

When studying the field emissivity of planar cathodes, particularly those made of graphite powders and by chemical vapor deposition, we repeatedly observed bright luminous rings on the screen. The rings are diffusely illuminated phosphor areas surrounding the image. Their diameter frequently exceeds the size of the image, and their brightness increases with the emission current.

Studies of the pulsed field emission process in an electron projector [1–5] revealed unambiguously that a bright ring around the image appears as the current density approaches the ultimate value. This effect was first described by Dyke *et al*. [1]. The rings were observed when the density of the emission current from a tungsten tip approached the ultimate value. The appearance of the rings is accompanied by a jump in the emission current. Dyke *et al*. explained this effect by thermal–field emission. The cathode is heated by the field emission current, and the heating makes an additional contribution to the current. In [2–5], predischarge phenomena, i.e., an anomalous increase in the emission current and the ring effect, were investigated with tips made of various materials. Several concentric rings were observed under certain conditions. Since the predischarge phenomena were sluggish and their duration depended on the time of passage of the field emission current, they were assumed to be of thermal nature. However, the nature of the rings was not conclusively established in the works cited because of the lack of experimental data. It was hypothesized that the rings are due to thermal–field emission from the circumference of the tip or asperities on the emitter surface and also that they are caused by electron diffraction. The first of the three hypotheses has received the most support in experiments and is considered the basic hypothesis today. Its primary disadvantage is that it fails to explain why the current in the ring increases by two or three orders of magnitude during the pulse while a temperature addition to the field-emission current (within the applicability of the Fowler–Nordheim equation) is only several tens of percent. Simulation performed in [6] showed that the rings may appear as a result of thermal–field emission from the circumference of the tip. However, this model cannot explain the appearance of several rings on the emission image.

MODEL OF RING FORMATION IN THE CASE OF PLANAR ANODE–CATHODE GEOMETRY

In [1–5], the rings were detected at the prebreakdown stage. Their occurrence was accompanied by a spontaneous growth of the current, causing the cathode breakdown and degradation. In our experiments, unlike [1–5], the rings persisted for a long time and their appearance did not cause the degradation of emitting centers. Figure 1a shows the time instant of ring origination. A further fourfold increase in the current adds substantially to the intensity of the rings, but the cathode remains operable (Fig. 1b).

Thus, the rings observed by us cannot be assigned to the prebreakdown stage, contrary to [1–5]. In the case of a planar field-emission cathode, the most plausible mechanism of ring formation appears to be the following. Primary electrons striking the anode knock out secondary electrons, which move to the cathode under the action of the electric field, forming a ring.

Consider a model of a planar diode with an electrode spacing L (Fig. 2). If the voltage applied to the anode is U, the field in the electrode gap is $E = U/L$. When considering the motion of secondary electrons near the anode, we ignore the existence of emitting centers on the cathode. They enhance the field near the cathode without affecting significantly the behavior of the electrons near the anode. Let a secondary electron with an initial energy ε_S leave the anode surface at an angle α.
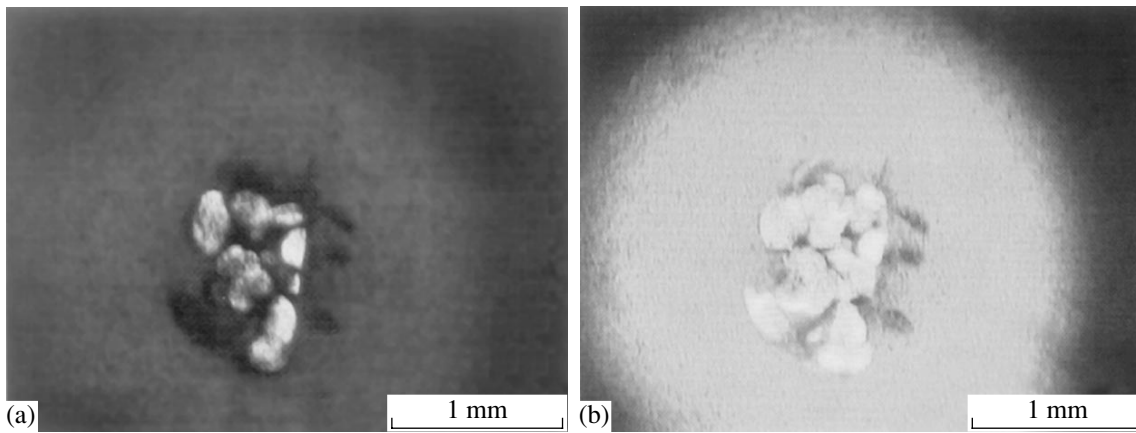
Fig. 1. Field-emission images of the rings. The anode–cathode voltage and current are (a) 2.84 kV and 10 μA and (b) 4.07 kV and 40 μA.

Straightforward analysis of the electron trajectory in a constant electric field allows us to find the distance of the electron to the secondary emission site in the anode plane:

$$r = 2\frac{\varepsilon_S}{eU}L\sin 2\alpha. \qquad (1)$$

The maximum distance to the anode is

$$H = \frac{\varepsilon_S}{eU}L(\sin\alpha)^2. \qquad (2)$$

Thus, the outer radius of the ring formed by secondary electrons with an initial energy ε_S is given by

$$r_{max} = 2L\frac{\varepsilon_S}{eU}. \qquad (3)$$

Assuming that elastically reflected electrons play a major role (specifically, they produce a ring of maximal radius), one may set $\varepsilon_S/eU \approx 1$. Hence, the maximal diameter of the ring is $D_{max} \approx 4L$.

Thus, the ring diameter depends only on the anode–cathode spacing and does not depend on the cathode current and anode–cathode voltage. Since elastically reflected electrons constitute only a fraction of the secondary electrons knocked out from the anode and since the secondary electrons escape the surface at different angles, the entire area between the outer boundary of the ring and the image will be "illuminated."

EXPERIMENTAL DATA

To verify our model, we tried to observe the rings in planar diode structures with variable electrode spacings. The spacing was adjusted with glass spacers of thickness 200, 400, and 600 μm. A glass plate with a transparent conducting coating and a phosphor layer served as an anode.

Figure 3 shows the image on the phosphor screen for an electrode spacing of 200 μm. Several bright rings, each formed around a single emitting center, are distinctly seen. All the rings are of the same diameter, although their contributions to the total emission current (i.e., their intensities) differ. Thus, the emission current does not influence the ring diameter.

The ring diameter versus electrode spacing is shown in Fig. 4. This dependence is given by formula (3). If the electrode spacing is kept constant, the ring diameter
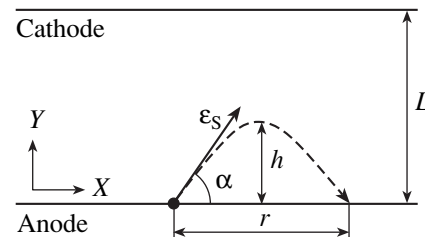

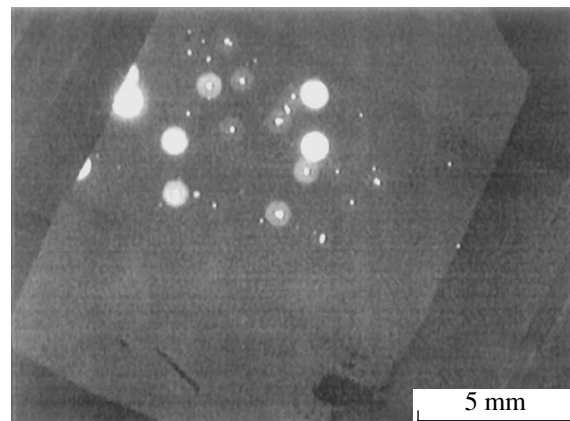**Fig. 2.** Model of ring formation in a planar diode.

Fig. 3. Field-emission image of several rings with different brightness (voltage 1.8 kV, total cathode current 80 μA).
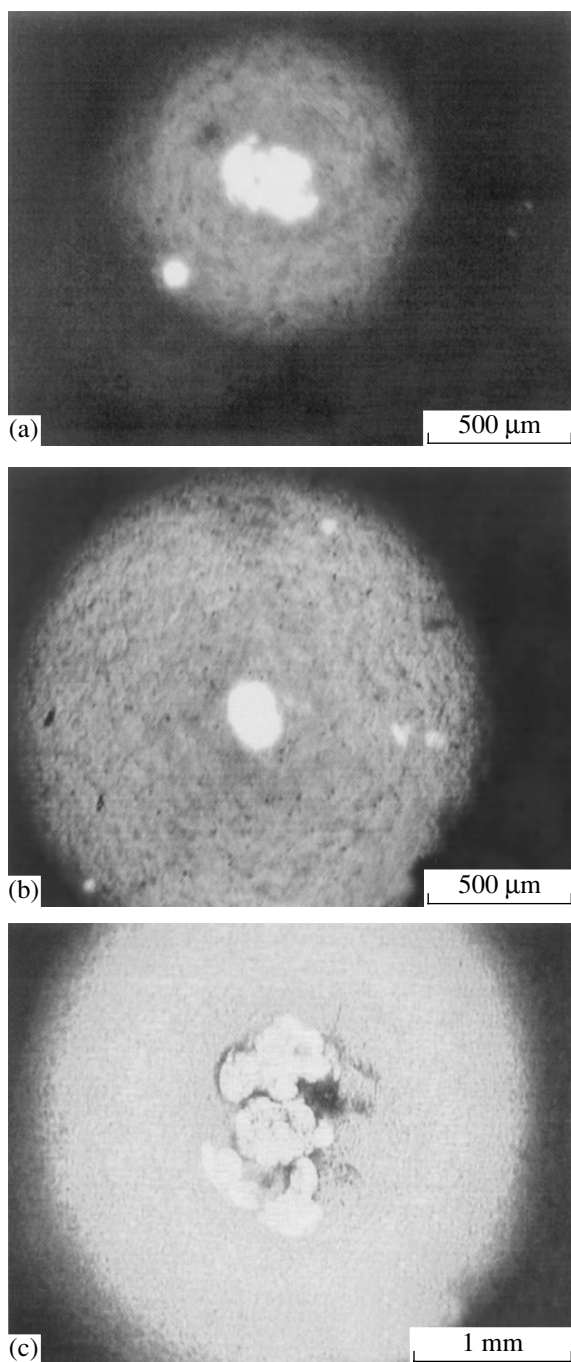
(a)   500 µm

(b)   500 µm

(c)   1 mm

Fig. 4. Field-emission images of the rings for an electrode separation of (a) 200, (b) 400, and (c) 600 µm.

Fig. 5. Position of two nearby emitting centers.
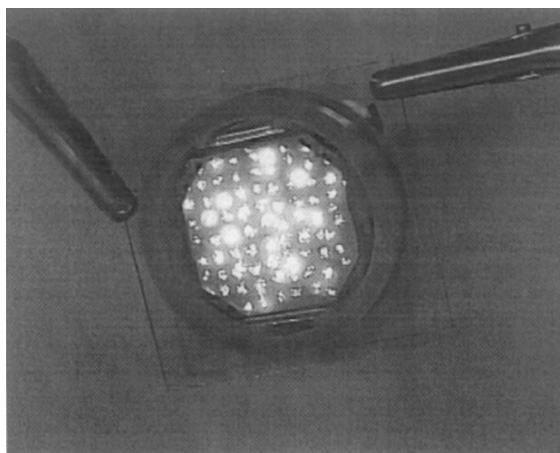
Fig. 6. Uniform-illumination lamp.

Fig. 7. Static glowing sign.

is independent of the operating voltage. Thus, it may be concluded that elastic reflection of electrons from the anode surface is the governing mechanism of ring formation.

To determine the critical current density at which the rings occur, we fabricated a planar diode with a cathode diameter of ≈150 µm. This value is *a fortiori* smaller than the ring size when the electrode spacing is 600 µm. For such a cathode, it is easy to find the critical
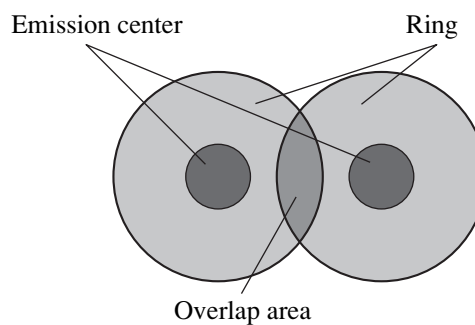
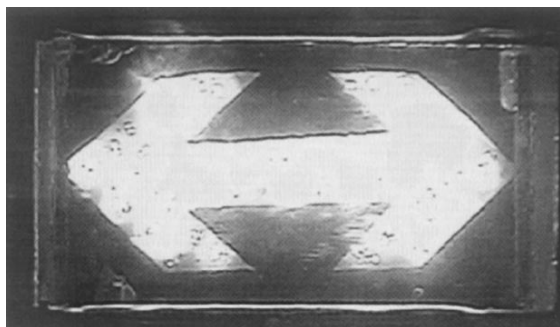current density (at which the rings appear) by measuring the total current. For the planar geometry to be preserved, the cathode was placed at the center of a metallic plate whose plane was parallel to the plane of the anode.

It was established that the ring nucleates at a current density in the primary spot on the order of 10^{-3} A/cm^2. The diameter of the ring observed did not depend on

either the cathode current or the anode–cathode voltage and was defined by formula (3).

In [1–5], the ring effect was detected in the case of spherical geometry. The anode was made in the form of a spherical flask covered by a conducting film and a phosphor layer, and a tip at the center of the sphere served as an emitter. Calculations carried out in those papers gave expressions similar to (1)–(3), and it turned out that the radius of a ring to be formed would approximately equal the diameter of the flask. This means that the effect that took place at low current densities conceptually cannot be observed under the experimental conditions used in [1–5].

APPLICATIONS

The ring formation described in this paper may have both a detrimental and a beneficial effect. For example, the ring may deteriorate the resolution of a flat-screen display but be useful in designing light sources with field-emission cathodes where uniform illumination of the phosphor must be provided.

Below, we describe a prototype light source utilizing the rings. The source consists of two parallel glass plates, which are substrates for the anode and the cathode. The anode substrate is covered by a conducting film and a phosphor layer. The electrode separation is fixed with glass spacers and equals 0.8 mm. The cathode is made of graphite foil adhered to the substrate. Emitting centers 2.5 mm distant from each other are arranged in a regular manner on the surface. The size of each of the centers is 1 mm. In this design, the diameter of the rings will be 3.2 mm and the ring around one center will not overlap a nearby center (Fig. 5).

In this situation, emitting centers will not be bombarded by residual gas ions being released from adjacent centers. At the same time, the rings will overlap each other, providing uniform illumination of the anode. The prototype structure was placed in a flask under a residual pressure of less than 10^{-6} torr, and field-emission tests were conducted. A photo of the operating device is shown in Fig. 6.

Ring illumination may also be helpful in producing static glowing signs, as exemplified in Fig. 7.

CONCLUSIONS

We constructed a model to account for the formation of rings on field-emission images that differs from that described in [1–6]. It considers processes taking place at the anode and is therefore applicable to devices with thermionic cathodes. It is found that (i) the diameter of the rings depends only on the electrode separation and (ii) the intensity of the rings varies with the primary current density and the coefficient of reflection of electrons from the anode coating.

The rings observed may be used in applications where anode uniform illumination is necessary. Examples are light sources with field-emission cathodes and static glowing signs.

REFERENCES

1. W. P. Dyke, J. K. Trolan, E. E. Martin, *et al.*, Phys. Rev. **91**, 1043 (1953).
2. I. L. Sokol'skaya and G. N. Furseĭ, Radiotekh. Élektron. (Moscow) **7**, 1474 (1962).
3. I. L. Sokol'skaya and G. N. Furseĭ, Radiotekh. Élektron. (Moscow) **7**, 1484 (1962).
4. G. N. Furseĭ and I. D. Tolkacheva, Radiotekh. Élektron. (Moscow) **8**, 1210 (1963).
5. G. N. Furseĭ, Zh. Tekh. Fiz. **34**, 1210 (1964) [Sov. Phys. Tech. Phys. **9**, 938 (1964)].
6. A. V. Batrakov and D. I. Proskurovskiĭ, Pis'ma Zh. Tekh. Fiz. **25** (11), 57 (1999) [Tech. Phys. Lett. **25**, 444 (1999)].

Translated by V. Isaakyan

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Oxygen Diffusion in Uranium Dioxide in the Temperature Range of Phase Transitions

**A. Ya. Kupryazhkin, A. N. Zhiganov, D. V. Risovanyĭ,
V. D. Risovanyĭ, and V. N. Golovanov**

Ural State Technical University, ul. Mira 19, Yekaterinburg, 620002 Russia
e-mail: kupr@dpt.ustu.ru
Received May 8, 2003

Abstract—The structure of and oxygen diffusion in UO_2 are studied by the molecular dynamics method in the range of transition to the superionic state (melting of the oxygen sublattice) and near the melting point of UO_2. The temperature dependence of the diffusion coefficient of a doubly charged oxygen ion in UO_2 is constructed. In the crystalline state at temperatures between 1800 and 2600 K, this dependence is described by an exponential dependence with a diffusion activation energy of 2.6 ± 0.2 eV. In the superionic state (2600–3100 K), the activation energy of diffusion of an oxygen anion decreases to 1.88 ± 0.13 eV. In melt (3100–3600 K), the exponential dependence of the diffusion coefficient of O^{2-} persists but the activation energy of diffusion decreases still further, to 0.8 ± 0.2 eV. Our experimental results agree (within the limits of experimental error) with data on oxygen diffusion in the crystalline phase obtained by other researchers. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

To predict the properties of uranium dioxide–based nuclear fuel under operating conditions and upon postradiation annealing, it is necessary to know the mass transfer coefficients for UO_2 constituent atoms in a wide temperature range. Diffusion of oxygen is of special interest because of its high mobility, which is associated with anti-Frenkel ordering in the UO_2 crystal lattice, and the superionic transition (melting of the anion sublattice) at temperatures of $\approx 0.8 T_m$ (where T_m is the melting point) [1]. Due to experimental difficulties, the data for oxygen diffusion [2] were obtained at relatively low temperatures. Therefore, they cannot be applied in the temperature range of the superionic phase and, especially, above the melting point.

Kurosaki *et al.* [3] showed that the properties of oxygen in UO_2 can be characterized by the molecular dynamics (MD) method. However, these researchers did not simulate the mobility and diffusion coefficient of oxygen. In [4], the diffusion coefficient was calculated for two temperatures in the superionic phase. However, the number of particles in the simulation was small, which makes it impossible to evaluate the potentialities of the MD method and the model proposed in [4], as well as the reliability of the experimental evidence. The issue of whether or not the ionic-crystal approach can be applied to describe mass transfer in UO_2 is also of fundamental importance.

In this work, we simulate oxygen diffusion by the MD method in a wide temperature range (covering the crystalline state, the superionic phase, and melt) and

derive the temperature dependences of the oxygen diffusion coefficient.

SIMULATION PROCEDURE

The crystal to be simulated consists of a basic crystallite incorporating N quadruply charged uranium ions and $2N$ doubly charged oxygen ions (a total of $3N$ particles), which constitute the UO_2 crystal lattice. In the calculation, we used periodic boundary conditions that are realized by translating the basic crystallite along three axes.

The interaction energy between the ions is represented as the sum of the non-Coulomb and Coulomb components:

$$U(t) = U_{nc} + U_c. \tag{1}$$

The non-Coulomb component is found by merely summing the interaction energies over the number of particles ($3N$) of the basic crystallite:

$$U_{nc}(t) = \frac{1}{2}\sum_{i=1}^{3N}\sum_{j \neq i}^{3N} \varphi(|\mathbf{r}_i(t) - \mathbf{r}_j(t)|), \tag{2}$$

where $\mathbf{r}_{i,j}$ are the radius vectors of crystallite particles,

$$\varphi_{\alpha\beta}(r_{ij}) = A_{\alpha\beta}\exp\left(-\frac{r_{ij}}{B_{\alpha\beta}}\right) - \frac{C_{\alpha\beta}}{r_{ij}^6}, \tag{3}$$

$A_{\alpha\beta}$ and $B_{\alpha\beta}$ are the repulsion constants, and $C_{\alpha\beta}$ is the dispersion interaction constant in the potential of interaction between particles of sorts α and β.

Table 1. Parameters of the interaction potentials for ions in uranium dioxide [4]

A_{++}, eV	A_{+-}, eV	A_{--}, eV	B_{++}, Å	B_{+-}, Å	B_{--}, Å	C_{++}, eV Å^{-6}	C_{+-}, eV Å^{-6}	C_{--}, eV Å^{-6}
–	1297.44	22764.3	–	0.3747	0.149	–	–	20.37
–	873.32735	50259.3	–	0.40369	0.15285	–	–	72.653

To simplify the calculation of the Coulomb interaction energy for particles with charges q_i and q_j, we applied the Ewald method [1] and reduced the expression for U_c to

$$U_c = \frac{1}{2}\sum_i \sum_{j \neq i} q_i q_j$$

$$\times \sum_{\mathbf{K}_m \neq 0} A(\mathbf{K}_m)\cos(\mathbf{K}_m(\mathbf{r}_i - \mathbf{r}_j)) + \sum_i \frac{q_i^2}{2}\sum_{\mathbf{K}_m \neq 0} A(\mathbf{K}_m) \quad (4)$$

$$+ \frac{1}{2}\sum_h \sum_i \sum_{j \neq i} q_i q_j \frac{\mathrm{erfc}(\varepsilon|\mathbf{r}_i - \mathbf{r}_j + \mathbf{h}|)}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{h}|} - \sum_i q_i^2 \frac{\varepsilon}{\sqrt{\pi}}.$$

Here, summation over i and j is performed for all particles of the system; summation over $\mathbf{h}$ is performed for all cells being translated ($\mathbf{h} = (k, l, m)$ with $-n_k \leq k, l, m \leq n_k$, where n_k is the number of neighboring cells being translated that are taken into account) except for the case where $\mathbf{h} = (0, 0, 0)$ and, at the same time, $i = j$; ε is the parameter controlling convergence in the Ewald method;

$$\mathbf{K}_m = (K_{mx}, K_{my}, K_{mz}) = \frac{2\pi}{L}(k, l, n)$$

is the reciprocal lattice vector ($k, l, n \subset [-\infty, \infty]$); and

$$A(\mathbf{K}_m) = \frac{4\pi}{(a_0 L)^3}\frac{\exp(-\mathbf{K}_m^2/4\varepsilon^2)}{\mathbf{K}_m^2},$$

where $(a_0 L)^3$ is the volume of a simulation cell, a_0 is the lattice parameter of UO$_2$, and L is the number of unit cells per edge of a cubic cell.

In deciding on a particular form of interaction potential, we considered potentials of two types given in [4] (Table 1). The first potential is derived in terms of the rigid-ion model, and the second is the modification of the first potential with regard for the results of simulation in terms of the shell model (ion polarization).

The potentials were tested by calculating the melting point in the course of simulating the basic crystallite (zero boundary conditions). With the first potential, the melting point of the crystallite was found to be $T_m = 3100 \pm 60$ K, which is close to the experimental value within the limits of experimental error. With the second potential, T_m was 300 K higher. Therefore, we used only the first potential in the simulation.

The equations of motion were simulated by the finite-difference method with a half-step. The compo-

nents of the coordinates, x_j^i, and velocities, v_j^i ($i = 1, 2, 3$), of a jth particle ($j = 1, 2, ..., N$) varied with time as

$$x_j^i(t + \Delta t) = x_j^i(t) + \Delta t\, v_j^i(t + \Delta t/2),$$
$$v_j^i(t + \Delta t/2) = v_j^i(t - \Delta t/2) + \Delta t m^{-1} F_j^i(t). \quad (5)$$

Here, $F_j^i(t)$ is the external force acting on the jth particle:

$$F_j^i(t) = \sum_{k, k \neq j} \partial\varphi(|\mathbf{r}_j(t) - \mathbf{r}_k(t)|)/\partial x_j^i. \quad (6)$$

In experiments, the number of particles in the basic crystallite was $3N = 1500$ and the time step, $\Delta t = 5 \times 10^{-15}$ s. The lattice parameter of the crystallite was taken from experimental data in [5].

With these parameters, the system being simulated comes to equilibrium in 100–200 steps.

RESULTS AND DISCUSSION

To analyze changes in the UO$_2$ structure upon heating, consider the pair radial distribution functions of ions in the crystal:

$$g_{\beta\beta}(r) = \frac{V}{N_\beta}\frac{N_\beta(r, \Delta r)}{4\pi r^2 \Delta r}. \quad (7)$$

Here, N_β is the number of particles β in a volume V and $N_\beta(r, \Delta r)$ is the number of particles in a spherical layer of thickness Δr located at a distance r from the central particle. The radial functions for oxygen, g_{OO}, and uranium, g_{UU}, for several temperatures are given in Fig. 1. At low temperatures up to $T_t = 2600$ K, the behavior of the radial functions corresponds to the crystalline state of the uranium and oxygen sublattices and the positions of the peaks correlate with certain coordination spheres. At 2600–3100 K, the behavior of the radial function of oxygen is virtually identical to its behavior in the melt at $T > 3100$ K. Although the lines of the radial function of uranium are broadened due to the intense motion of uranium cations, the positions of the peaks remain virtually unchanged, which corresponds to the crystalline state of the uranium sublattice. At temperatures above 3100 K, the radial functions of uranium and oxygen correspond to the melt.

Thus, the results of structural study indicate that the ionic model predicts adequately the superionic transition (i.e., melting of the oxygen sublattice) at $T_t = $
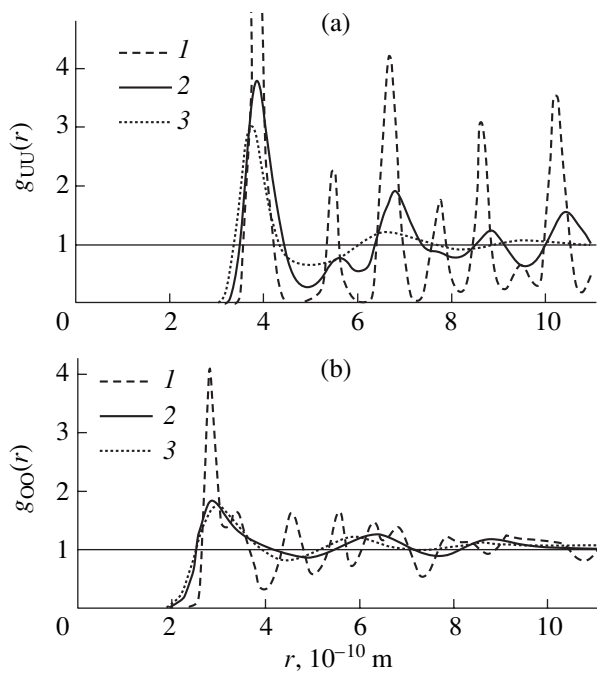
Fig. 1. Radial distribution functions for (a) uranium ions and (b) oxygen ions in UO_2 at (*1*) 300, (*2*) 2900, and (*3*) 3200 K.

Fig. 2. Time dependence of the MQDs of (*1*) oxygen ions and (*2*) uranium ions for (a) $T = 2900$ and (b) 3400 K.

2600 K and the melting point of the crystal ($T_m = 3100 \pm 60$ K). The latter value coincides with experimental findings [2] within the experimental accuracy.

To determine the diffusion coefficients of oxygen anions at various temperatures, we calculate the mean quadratic displacements (MQDs) of the anions by the relationship

$$\langle \Delta r^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{r}_i(t) - \mathbf{r}_i(0)|. \tag{8}$$

Typical MQD curves for the oxygen and uranium ions are shown in Fig. 2.

As follows from the calculation and Fig. 2, $\langle \Delta r^2 \rangle$ for oxygen exhibits well-defined linear variation with time throughout the t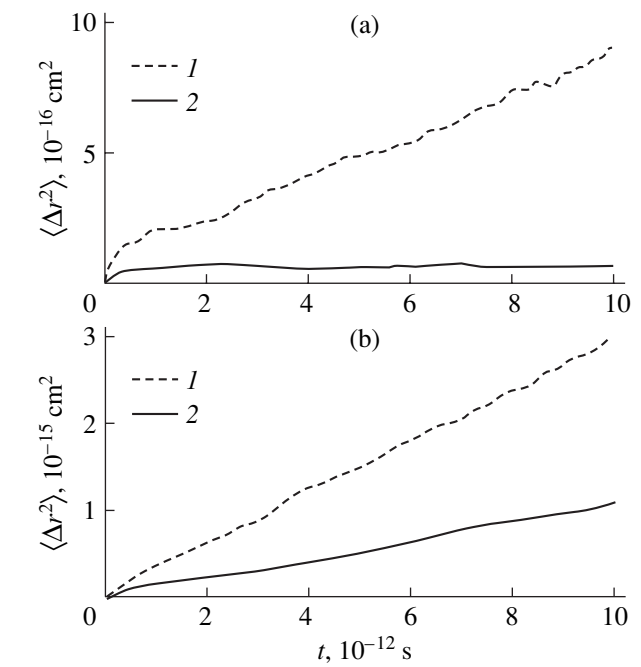emperature range. For uranium ions, a distinct slope is observed only at high temperatures above the melting point of the crystal.

From the time dependence of the MQDs, which is fitted by the formula

$$\lim_{t \to \infty} \langle \Delta r^2 \rangle = \text{const} + 6Dt, \tag{9}$$

we determine the diffusion coefficients of oxygen in a wide temperature range (1000–3600 K). The calculation results are given in Fig. 3. By processing the temperature dependences in Fig. 3, one can distinguish three ranges, corresponding to the (I) crystalline, (II) superionic, and (III) molten states. In each of the ranges, the diffusion coefficient of oxygen in uranium dioxide depends on temperature as

$$D = D_0 \exp\{-E^d/kT\}. \tag{10}$$

The preexponentials D_0 and the activation energies of diffusion E^d for the three ranges are listed in Table 2, which also contains the experimental data obtained in [2]. The temperature dependences of the diffusion coefficient and the results of simulation by the MD method [4] are shown in Fig. 3.

As follows from Table 2 and Fig. 3, on the transition from the crystalline to the superionic state, the activation energy of diffusion of oxygen ions decreases by a factor of 1.4 and the preexponential drops roughly 30-fold, with the diffusion coefficient experiencing no jump (within the limits of experimental error).

Table 2. Parameters of the diffusion coefficients for oxygen ions in UO_2

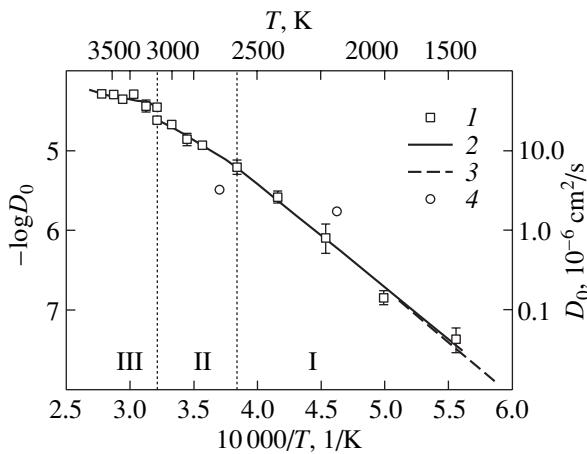| Temperature	D_0, cm^2/s	E^d, eV	Phase state
1800–2600	$(0.8^{+2.9}_{0.2})$	2.6 ± 0.2	Crystal
2600–3100	$(2.6^{+1.9}_{-1.1}) \times 10^{-2}$	1.88 ± 0.13	Superionic state
3100–3600	$(6.6^{+7.6}_{-3.5}) \times 10^{-4}$	0.8 ± 0.2	Melt
1100–1500	1.44	2.74	Crystal [2]

Fig. 3. Temperature dependence of the diffusion coefficients of oxygen ions near the phase transitions in uranium dioxide: (*1*) calculated diffusion coefficients of oxygen, (*2*) approximation, (*3*) extrapolation of the experimental data for the diffusion coefficients of oxygen [2], and (*4*) data from [4].

cies E^{f}, and the activation energy of diffusion of vacancies E^{v} are related as

$$E^{\mathrm{d}} = \frac{1}{2}E^{\mathrm{f}} + E^{\mathrm{v}}. \tag{11}$$

If we take the values $\overline{E}^{\mathrm{f}} \approx 5.3$ eV and $\overline{E}^{\mathrm{v}} \approx 0.3$ eV, averaged for different potential types with allowance for ion polarization [4], then the activation energy of diffusion $E^{\mathrm{d}} = 2.35$ eV obtained from (11) coincides with our experimental data, $E^{\mathrm{d}} = 2.6 \pm 0.2$ eV, within the limits of experimental error (Table 2).

As follows from Table 2 and Fig. 3, our results coincide with the experimental data for the crystalline state within the limits of experimental error. This fact supports both the validity of the method and the efficiency of the ionic model applied. The discrepancy between the results of simulation in [4] and in this work (Fig. 3) is likely to be associated with the small number of particles (96 ions) used in [4] and a short time of computation (unfortunately, this value in [4] is omitted).

Once the crystallite has melted completely, the diffusion coefficient increases stepwise by a factor of almost 1.5 at the melting point, obviously because of a sharp decline (by $\approx 9.6\%$) in the uranium dioxide density upon melting. The activation energy of diffusion of the ions in the melt is lower than in the superionic state by a factor of 2.4. The relatively high activation energy of oxygen anion diffusion in the melt (0.8 eV) is likely due to the high charges of the oxygen and uranium ions.

The results obtained for the crystalline state agree with the vacancy mechanism of oxygen ion diffusion in uranium dioxide, which shows anti-Frenkel disorder of the oxygen sublattice. In this case, the activation energy of diffusion E^{d}, the energy of formation of anion vacan-

REFERENCES

1. L. V. Matveev and M. S. Veshunov, Zh. Éksp. Teor. Fiz. **111**, 585 (1997) [JETP **84**, 322 (1997)].

2. Hj. Matzke, in *Nonstoichiometric Oxides*, Ed. by O. Toff Sorensen (Academic, New York, 1981), pp. 155–232.

3. K. Kurosaki, K. Yamada, M. Uno, *et al.*, J. Nucl. Mater. **294**, 160 (2001).

4. J. R. Walker and C. R. A. Catlow, J. Phys. C **14**, L979 (1981).

5. D. M. Skorov, Yu. F. Bychkov, and A. I. Dashkovskiĭ, *Reactor Materials Science* (Atomizdat, Moscow, 1979).

Translated by K. Shakhlevich

=================== EXPERIMENTAL INSTRUMENTS ===================
AND TECHNIQUES

Fullerene Films Highly Resistant to Laser Radiation

M. A. Khodorkovskii*, S. V. Murashov*, T. O. Artamonova*, A. L. Shakhmin*,
A. A. Belyaeva*, and V. Yu. Davydov**

* Prikladnaya Khimiya Russian Research Center, St. Petersburg, 197198 Russia
e-mail: mkhodorkovskii@rscac.spb.ru
** Ioffe Physicotechnical Institute, Russian Academy of Sciences,
Politekhnicheskaya ul. 26, St. Petersburg, 194021, Russia
Received May 7, 2003; in final form, July 17, 2003

Abstract—The effect of laser radiation on the structure and phase composition of thin fullerene films is studied. Raman measurements show that fullerene films applied with a supersonic molecular beam remain structurally stable even if the laser power density is many hundreds of times higher than that at which thermally deposited films degrade. A plausible reason for a high laser radiation hardness of the films is fast polymerization with the formation of linear and multidimensional configurations. © 2004 MAIK "Nauka/Interperiodica".

INTRODUCTION

The discovery of fullerenes has given impetus to studying the interaction between laser radiation and fullerene films applied on different coatings. According to a number of parameters, fullerene is one of the best photoconducting organic materials [1]. It is also viewed as a candidate material for frequency doubling and tripling optical converters [2]. Moreover, fullerene was tested as a material for optical switches [3]. Thus, study of laser radiation–fullerene film interaction and the search for new technologies for radiation-resistant fullerene coatings are of current interest.

In [4], we suggested a method of applying thin films with a pulsed supersonic molecular beam (SMB) seeded by a material to be deposited and compared the potentialities of the new method with those of conventional thermal deposition. Based on unique properties of SMB fullerene films, it was concluded that the material is promising for many applications. It was also noted that SMB fullerene films retain their molecular constitution when exposed to very high laser radiation loads (i.e., offer a high laser radiation hardness), which makes them still more promising coating materials. In this work, we study the properties of fullerene films obtained by various methods to reveal reasons for the high radiation hardness of SMB films.

The properties of fullerenes are usually examined with Raman spectroscopy. Here, the effect of the environment on the molecules under study is studied in the most sensitive range of the Raman spectrum (the peak at 1460–1470 cm^{-1}), which corresponds to the tangential Ag(2) mode of fullerene pentagon vibration (pentagonal pinch mode) [5, 6]. For an unpolymerized (pristine) C_{60} film applied by conventional thermal deposition, the position of this peak is 1469 cm^{-1}. UV irradiation of this film causes the photopolymerization

(dimerization) of fullerene and shifts the peak to 1459 cm^{-1} [7]. Similar behavior of the Raman spectrum is observed when the laser irradiation dose increases in the visible range (488 nm) [8]. It was speculated [8–10] that the reaction of photopolymerization breaks double bonds in pentagonal segments and forms a bridge between two C_{60} molecules. A high sensitivity of the Raman spectroscopy method in this spectral range makes it possible to study other high-pressure high-temperature polymer phases of fullerene (HPHT polymers), such as orthorhombic (O), tetragonal (T), and rhombohedral (R) polymers [11, l2]. In this work, we study the general properties of fullerene and molecular bonding in it using high-sensitivity Raman spectroscopy.

EXPERIMENT AND DATA PROCESSING

Raman spectra from fullerene films were taken at room temperature with a 488-nm argon laser. The laser radiation was used to take Raman spectra and also to modify the phase composition of the film. The radiation power density was varied between 15 mW/mm^2 and 400 W/mm^2; the exposure time, from 2 min to 1 h. With these parameters varying in the above ranges, we were able to trace the fine structure of the Ag(2) line in a wide range of radiation doses. The spurious signal due to luminescence from the substrate was suppressed by applying the fullerene films on single-crystal magnesium oxide, making no contribution to the spectrum. The films were deposited using high-purity (99.92%) C_{60} as a source. The SMB deposition method (for details, see [4]) consists in irradiating the substrate surface by a supersonic beam of helium molecules that is seeded by fullerene molecules. In this method, the transport velocity v_{tr} of fullerene molecules is 371 m/s; their thermal velocity, $v_t = 112$ m/s; their temperature,

T = 560 K; and the Mach number, 4 [13]. Accordingly, the kinetic energy of fullerene molecules equals 0.526 eV, which is one order of magnitude higher than upon thermal deposition.

To estimate the radiation-induced modification of the film's phase composition, the resulting Raman spectra were resolved into components with a specially designed program that takes into account the position and shape of Raman spectra taken from different polymers. The percentage of the phases polymerized was calculated from the areas under respective components. When designing the program, we used spectroscopic data in [12]. The shape of the Raman spectrum for each polymer was approximated by a set of six to seven simple Lorentzian functions. The shape of the Ag(2) peak for pristine C_{60} was represented by the mixed Gaussian–Lorentzian function [14] with the Hg(7) satellite at 1423 cm^{-1}.

It should be noted that recent publications [15, 16] report Raman spectra for various polymeric forms that differ from those described in [7, 12]. Without going into the discussion among the authors of the works cited, we compare our data with the classical results obtained in [12].

RESULTS

Figure 1 shows the Raman spectrum near the Ag(2) mode vs. the dose of irradiation by the SMB. The radiation is seen to modify the initial spectrum and lead to additional features. These features indicate that the phase composition becomes more complex than the initial one. To identify the new composition, the spectra obtained were resolved into components corresponding to unpolymerized fullerene and its polymerized fractions, as described above. As an example, Fig. 2 demonstrates the resolved spectrum taken at a dose of 40 J/mm^2. At low laser intensities, only two fullerene phases, pristine C_{60} and O polymer (fundamental vibration frequency 1463 cm^{-1} [12]), are present. Initially, the film contains about 60% pristine C_{60}; then, its amount drops to 10% as the irradiation dose grows. A close amount of pristine C_{60} is found when the Hg(8) line is resolved (Fig. 3). Its peak shifts from 1575 cm^{-1} for the unpolymerized phase to 1565 cm^{-1}, which is typical of fullerene fractions polymerized by irradiation [12]. One should bear in mind that even the low doses introduced into the target to take Raman spectra may cause partial polymerization of the fullerene in the as-deposited SMB films. To estimate this effect, we, based on the resolution of the Ag(2) peak, constructed the dependence of the phase composition of the SMB film on the irradiation dose (Fig. 4). As follows from Fig. 4, the radiation-induced linear polymerization of the fullerene is described well by an exponential whose value at the zero dose indicates that the as-deposited SMB film almost entirely consists of pristine C_{60}.
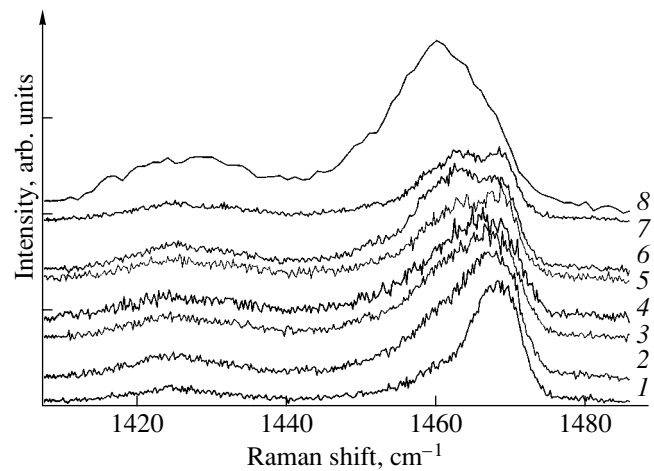

Fig. 1. Variation of the Raman spectra in the vicinity of the Ag(2) mode of fullerenes when SMB films are irradiated with a dose of (*1*) 2.9, (*2*) 5.82, (*3*) 8.76, (*4*) 11.7, (*5*) 14.6, (*6*) 25.1, (*7*) 40.0, and (*8*) 83 J/mm^2.
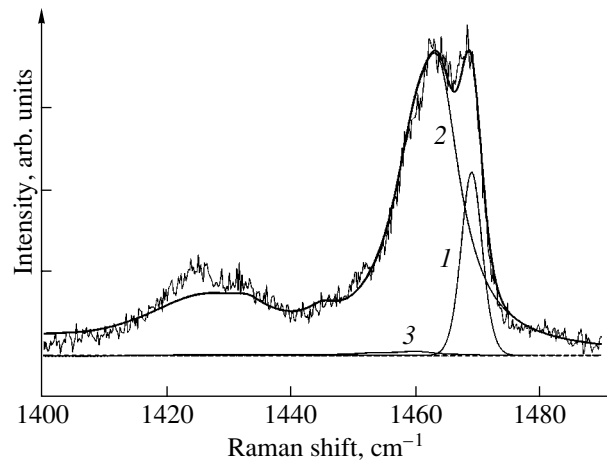

Fig. 2. Resolution of the Raman spectra taken at a dose of 40.0 J/mm^2. (*1*) Pristine polymer, (*2*) O polymer, and (*3*) photopolymer.

Thus, at radiation doses of no more than 40 J/mm^2, the film consists largely of the O polymer phase and pristine C_{60}, their percentage ratio varying with the dose. At higher doses, the resolution of the Ag(2) mode shows other polymer fractions. Figure 5 demonstrates the Raman spectrum obtained at a total dose of 82 kJ/mm^2 (the power density equals 86 W/mm^2 in this case). Here, the percentage ratio of the polymer fractions differs substantially from that at low doses: pristine C_{60}, ≈3%; photopolymers (dimers), ≈45%; O polymers, ≈36%; and R polymers, ≈16%. T polymers are absent.

Our composition versus dose dependences for SMB fullerene films are distinct from those obtained for thermally deposited films. As follows from publications, the latter contain only the photopolymer phase at low
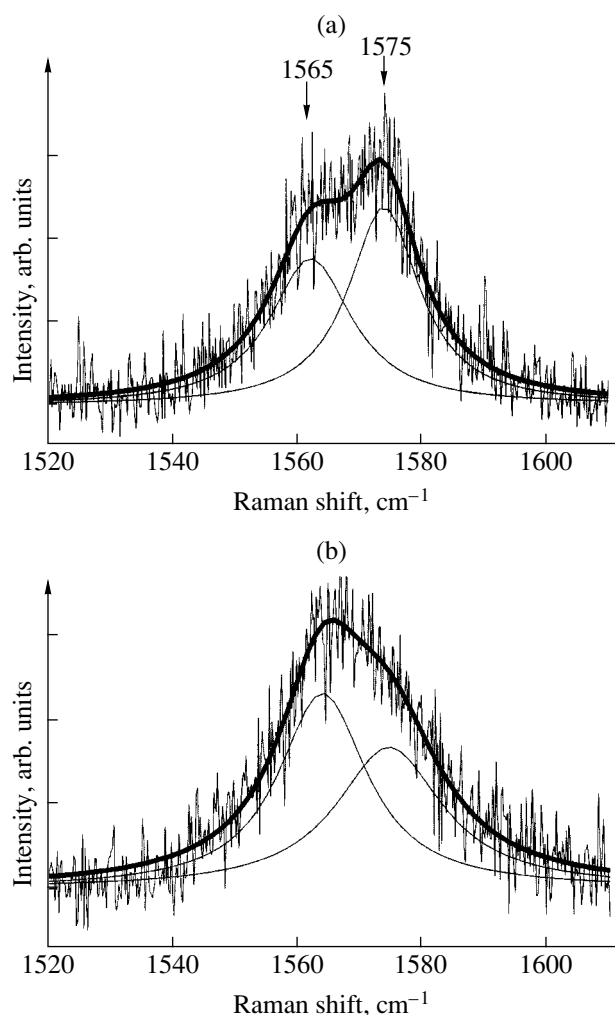
Fig. 3. Raman spectra near the Hg(8) line (a) at the beginning and (b) at the end of irradiation.
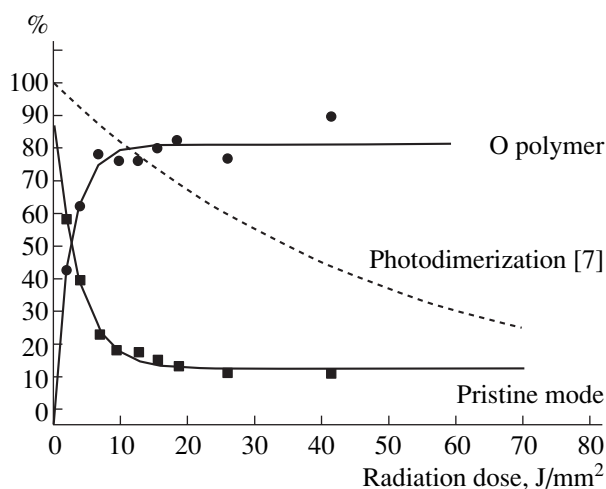

Fig. 4. Variation of the SMB film phase composition with irradiation dose (the graph is based on the resolution of the Ag(2) peak) and the result of polymerization of thermal fullerene [7].

doses. A rise in the dose leads to the complete degradation of the molecular constitution of fullerene. As follows from Fig. 6, the molecular constitution of thermally deposited films degrades even at a laser power density of 4 W/mm², as evidenced by the disappearance of the Ag(2) mode and the appearance of broad peaks typical of graphite structures and disordered carboniferous inclusions [17] (curve *4*). Figure 6 also shows the Raman spectrum taken of the SMB film, which indicates that the molecular structure of the fullerene persists at much higher doses (80 W/mm², curve *2*). Moreover, even at a power density of 400 W/mm², when the SMB film starts evaporating, the Raman spectrum does not exhibit changes characteristic of degraded fullerene molecules (curve *3*). Such a high stability of fullerene films against laser radiation is observed for the first time. Earlier, radiation-resistant fullerene crystals synthesized by a special technology retained their molecular constitution up to 20 W/mm² (2000 W/cm²) [18].

DISCUSSION

As was noted, the formation of O, R, and T polymers takes place at high temperature and/or pressure. According to [12], O polymers are synthesized at a high pressure (4.8 GPa) and relatively low (for the HPHT technology) temperature (250°C). R polymers are formed at the same pressure and much higher temperatures (up to 700°C). Pure T polymer is difficult to produce. A mixed phase of O (50%) and T (50%) polymers was obtained at 1.1 GPa and 600°C in [12]. The least center-to-center spacing of fullerene molecules is 9.26, 9.2, and 9.09 Å for O, R, and T polymers, respectively [12]. Bulk polymers, which are produced at very high pressure (8.5 GPa) and temperature (750°C), have also been mentioned [19]. However, they are practically unfeasible; anyway, we are unfamiliar with publications concerning Raman spectra for this material. In photopolymers, which are, in essence, C_{60} dimers, the center-to-center distance of the molecules is 9.76 Å [12]. Photopolymers are produced by exposing thermal C_{60} films to UV or laser radiation.

As follows from the above, the irradiation of SMB films produces O polymers rather than photopolymers. To explain this fact, it is necessary to consider the formation of SMB films in greater detail.

When constructing a model of interaction between fast C_{60} molecules from an SMB and the surface, we can rely upon two works [20, 21] where the energy of C_{60}–surface and C_{60}–C_{60} bonds as a function of surface–molecule or molecule–molecule spacing is estimated.

Interaction of an adsorbed C_{60} molecule with a solid surface is governed by orientation, induction, and dispersion forces. The relationship between these forces, which controls the position (adsorption site) of the molecule, its orientation relative to the surface, and the adsorption energy, depends considerably on the sub-

strate type. The effect of the surface potential relief on the position of the molecule adsorbed depends on both the initial surface condition and the amount of the adsorbate. After the deposition of several monolayers, the effect of the substrate weakens and the structure of subsequent layers becomes independent of the structure of the interface.

It may be assumed that, as the kinetic energy of a molecule whose momentum is directed normally to the surface grows, the forces mentioned above and, hence, the potential field of the surface will affect the trajectory of the molecule near the surface to a lesser extent. The higher the kinetic energy of a C_{60} molecule, the weaker the dependence of its adsorption site, orientation, and bond energy with the surface on the surface properties. In this case, the fill of the interface and subsequent layers with adsorbed molecules will increase with their kinetic energy, with the packing becoming progressively closer. A high density of SMB films changes the π-electron collective oscillations, which are a result of interaction between electron shells of nearby C_{60} molecules [22]. This reduces the possibility of exciting π plasmons, as demonstrated by the characteristic energy loss spectra [22].

Thus, the SMB technology makes it possible to simulate the formation of fullerene structures under high pressure. However, C_{60} molecules in an SMB have a low temperature (≈ 550 K), so that double bonds do not break when the molecules are deposited on the surface and the polymers do not form. This process can be stimulated by any extra external action that raises the number of degrees of freedom of C–C bonds and favors their breaking with the formation of 2 + 2 double bonds with the nearest fullerene molecule. That is why even low-intensity laser irradiation of SMB films causes intense polymerization of C_{60} molecules and the formation of O polymer chains. The peak intensity of unpolymerized fullerene (I_{1469}) versus the irradiation dose D can be represented in the form

$$I_{1469}(D) = I_0 \exp(-D/k),$$

where I_0 is the initial fraction of unpolymerized fullerene and $k \approx 46.3$ is the parameter characterizing the activation energy of polymerization.

A similar dependence was obtained in [7] for the photoconversion of thermal fullerene films, with a shift of the peak from 1469 to 1459 cm^{-1}, which is typical of photopolymers (Fig. 4). The intensities of the peaks varied with the dose by the same exponential law, but k exceeded the value for SMB films 13 times. This means that the activation energy is low and, hence, polymerization will proceed rapidly upon weak external actions. This property of SMB films may find application in data-writing devices where highly localized low-intensity optical radiation is used, such as near-field optical microscopes [23, 24], which offer a resolution of about several hundred angstroms.
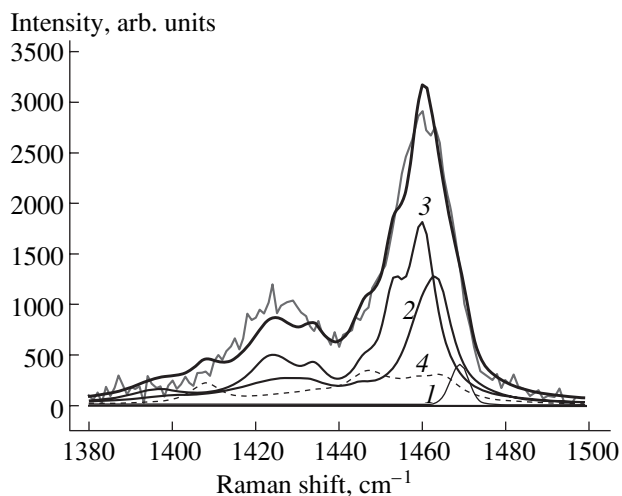
Fig. 5. Resolution of the Raman spectrum taken at a high irradiation dose (82.0 kJ/mm^2). (*1*) Pristine polymer, (*2*) O polymer, (*3*) photopolymer, and (*4*) R polymer.
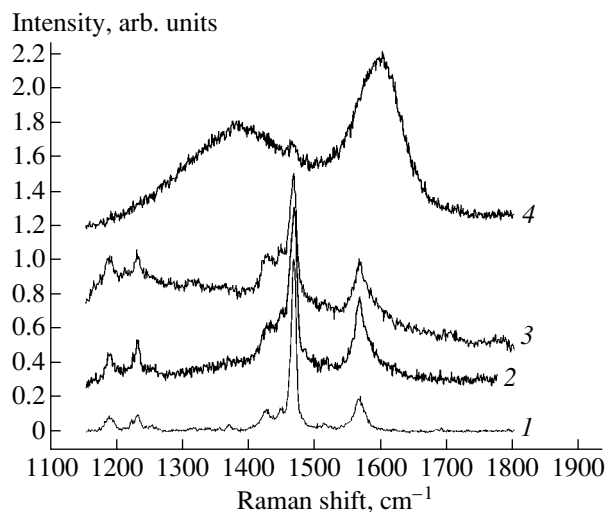

Fig. 6. Variation of the Raman spectra near the Ag(2) mode of (86%C_{60} + 14%C_{70}) fullerene. (*1*) Initial spectrum, (*2*) SMB film (laser power density $P = 80$ W/mm^2), (*3*) SMB film ($P = 400$ W/mm^2), and (*4*) thermal film ($P = 4$ W/mm^2).

Radiation-induced polymerization in SMB films may be responsible for their high radiation hardness. The radiation energy absorbed by a C_{60} molecule is transferred to the entire molecular ensemble of the polymer via bridges, so that the entire polymer accumulates the energy. This energy causes breaking of weaker bonds, which are polymer bonds rather than C–C bonds in a C_{60} molecule. In addition, the polymer structure imparts a part of the energy to the substrate.

In comparison with similar processes taking place in thermal films, the characteristic time of energy accumulation per C_{60} molecule in a SMB film changes appreciably and the rate of energy transfer through the more perfect interface increases.

A further rise in the laser radiation intensity apparently causes not only the rupture of linear polymerized chains with the formation of photopolymers (dimers) but also the production of HPHT polymers (R polymers), which are synthesized at temperatures much higher than the synthesis temperature of O polymers (700°C) [12].

We stress once more that the Raman spectra for the different polymer fractions have not been fully understood yet and are the subject of extensive discussion today. For example, Senyavin *et al.* [15] reported Raman spectra for O, R, and T polymers that differ greatly from the corresponding spectra in [12] and argued that their HPHT polymers are of much higher quality than those studied in [12]. A peak at 1463 cm^{-1} was assigned in [12] to the O polymer phase with double $2 + 2$ bonds. This phase is believed to be rather pure, unlike the T phase. The synthesis of the latter is considered to be difficult; therefore, the authors of [12] studied a T + R polymer mixture. In [15], the vibrational mode at 1463 cm^{-1} is related to polymerization with the formation of single C–C bonds between fullerene molecules (such bonds are present, e.g., in T polymers). Our idea is that fast polymerization of SMB films may provide a high resistance to laser radiation. The polymer fraction in this case may be arbitrary, including that suggested in [12].

CONCLUSIONS

Our results indicate that fullerene films synthesized with a pulse supersonic molecular beam contain the denser unpolymerized (pristine) fullerene fraction, which, when exposed to low-intensity laser radiation, passes into the polymer fraction typical of HPHT fullerenes. Polymer chains make it possible to distribute the absorbed energy over the entire ensemble (including the substrate). As a result, the molecular constitution of the fullerene framework persists at power densities far exceeding the critical power density for conventional fullerene coatings.

ACKNOWLEDGMENTS

REFERENCES

1. A. V. Eletskiĭ and B. M. Smirnov, Usp. Fiz. Nauk **165**, 977 (1995) [Phys. Usp. **38**, 935 (1995)].

2. H. Hoshi, N. Nakamura, Y. Maruyama, *et al.*, Jpn. J. Appl. Phys. **30**, L1397 (1991).

3. A. Kost, L. Tutt, M. Klein, *et al.*, Opt. Lett. **18**, 334 (1993).

4. M. A. Khodorkovskiĭ, A. L. Shakhmin, S. V. Murashov, *et al.*, Pis'ma Zh. Tekh. Fiz. **24** (10), 20 (1998) [Tech. Phys. Lett. **24**, 379 (1998)].

5. Zhou Ping, Dong Zheng-Hong, A. M. Rao, *et al.*, Chem. Phys. Lett. **211**, 337 (1993).

6. Wang Ying, J. M. Holden, P. C. Eklund, *et al.*, Chem. Phys. Lett. **217**, 413 (1994).

7. Wang Ying, J. M. Holden, Dong Zheng-Hong, *et al.*, Chem. Phys. Lett. **211**, 341 (1993).

8. A. M. Rao, Zhou Ping, Wang Kai-An, *et al.*, Science **259**, 955 (1993).

9. L. Akselrod, H. B. Byrne, C. Tomsen, *et al.*, Chem. Phys. Lett. **215**, 131 (1993).

10. T. L. Makarova, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 257 (2001) [Semiconductors **35**, 243 (2001)].

11. M. Nunez-Regueiro, L. Marques, J.-L. Hodeau, *et al.*, Phys. Rev. Lett. **74**, 278 (1995).

12. A. M. Rao, P. C. Eklund, J.-L. Hodeau, *et al.,* Phys. Rev. B **55**, 4766 (1997).

13. M. A. Khodorkovskiĭ, S. V. Murashov, T. O. Artamonova, *et al.*, Zh. Tekh. Fiz. **73** (5), 1 (2003) [Tech. Phys. **48**, 523 (2003)].

14. D. Briggs and M. P. Seah, *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (Wiley, New York, 1983; Mir, Moscow, 1987).

15. V. N. Senyavin, V. A. Davydov, L. S. Kashevarova, *et al.*, Chem. Phys. Lett. **313**, 421 (1999).

16. T. Wagberg, P. Jacobsson, and B. Sundqvist, Phys. Rev. B **60**, 4535 (1999).

17. S. M. Mominuzzaman, K. M. Krishna, T. Soga, *et al.*, Carbon **38**, 127 (2000).

18. M. Haluska, H. Kuzmany, M. Vybornov, *et al.*, Appl. Phys. A **56**, 161 (1993).

19. S. Itoch, P. Ordejon, D. A. Drabold, *et al.*, Phys. Rev. B **53**, 2132 (1996).

20. P. A. Gravil, M. Devil, Ph. Lambin, *et al.*, Phys. Rev. B **53**, 1622 (1996).

21. Gh. Girard, Ph. Lambin, A. Dereux, *et al.*, Phys. Rev. B **49**, 11425 (1994).

22. A. L. Shakhmin, A. M. Khodorkovskiĭ, S. V. Murashov, *et al.*, Pis'ma Zh. Tekh. Fiz. **27** (3), 1 (2001) [Tech. Phys. Lett. **27**, 87 (2001)].

23. S. Wegscheider, A. Krisch, J. Mlynek, *et al.*, Thin Solid Films **264**, 264 (1995).

24. T. Shintani, K. Nakamura, S. Hosaka, *et al.*, Ultramicroscopy **61**, 285 (1995).

Translated by V. Isaakyan

SHORT
COMMUNICATIONS

Hysteresis in the Static Aerodynamic Characteristics
of a Curved-Profile Wing

I. V. Kolin, V. G. Markov, T. I. Trifonova, and D. V. Shukhovtsov

Joukowski Central Aerohydrodynamic Institute, Zhukovskiĭ, Moscow Oblast, 140180 Russia

Received April 29, 2003

Abstract—The effect of profile asymmetry, Reynolds number, and angles of deflection of high-lift devices on the hysteresis loop shape in the dependences of the static aerodynamic forces and moments on angles of attack is studied. These dependences are measured in wind tunnels during testing of rectangular wings of equal thickness but different curvature. The wings have deflectable flaps spanning the wing trailing edge. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It was noted [1–14] that the dependences of aerodynamic forces and moments on the angle of attack for rectangular wings with aspect ratios $\lambda = 1$ and 5, as well as for aircraft with high-aspect-ratio straight wings, exhibit hysteresis at $Re \leq 4 \times 10^6$. The shape of the hysteresis loop boundaries was shown to depend on the Re [3], profile thickness [4], aspect ratio [5], wing surface roughness [13], and slip angles [5]. Based on visualization data, the flow structures under the test conditions corresponding to the outer and inner boundaries of multiple hysteresis were analyzed [1, 5].

In this work, emphasis is on the effect of the rectangular wing profile curvature on the hysteresis loop boundaries.

TEST CONDITIONS AND MODEL

Aerodynamic forces and moments were measured with six-component mechanical balances during testing of two models of rectangular wings in a low-subsonic wind tunnel. The geometries of the wings were as follows. Model 1: aspect ratio $\lambda = 2.25$; NACA 23010 asymmetric profile; and flap chord $b_f = 0.28b$, where b is the wing chord. Model 2: aspect ratio $\lambda = 2.0$, NACA 0010 symmetric profile, and flap chord $b_f = 0.4b$. The measurements were made in the moving coordinate system. The arrangement of the dynamic unit and model wings in the working part of the wind tunnel were described elsewhere [13].

TEST RESULTS

Typical dependences of the coefficients $c_y(\alpha)$ and $m_z(\alpha)$ for wings 1 and 2 with the undeflected flap ($\delta = 0$) at $Re = 2.02 \times 10^6$ when the angle of attack $\dot\alpha$ increases ($\dot\alpha > 0$) and decreases ($\dot\alpha < 0$) are shown in Fig. 1. It is seen that the curves exhibit hysteresis at $\alpha =$
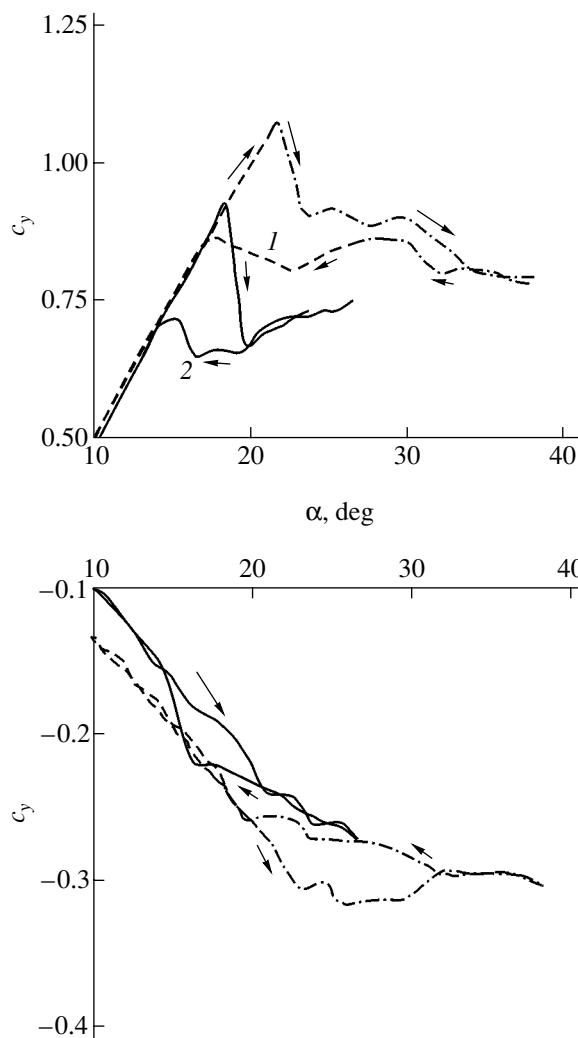

Fig. 1. Normal-force and longitudinal-moment coefficients vs. the angle of attack ($\delta = 0$) at $Re = 2.02 \times 10^6$ for (*1*) an asymmetric- and (*2*) a symmetric-profile wing (wings 1 and 2, respectively).
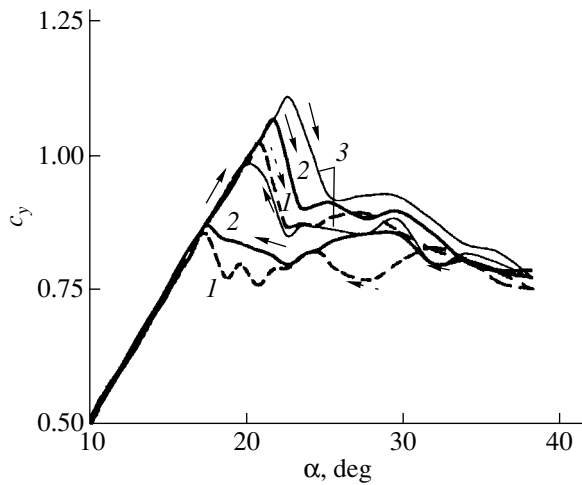
Fig. 2. Coefficient c_y for wing 1 at Re = (*1*) 1.64×10^6, (*2*) 2.02×10^6, and (*3*) 2.5×10^6.

Fig. 3. Polar of wing 1 for various c_{xa} and Re numbers (for *1–3*, see Fig. 2).

$17.5°–34°$ (wing 1) and $14°–20°$ (wing 2). The loop area for wing 1 is roughly twice that for wing 2. The difference between the maximal values of c_y is $\Delta c_{y\,max} = 0.15$. The critical angle of attack at which the transition from the upper to the lower branch of the curve $c_y(\alpha)$ takes place increases from $\alpha_{c1} = 18.5°$ for wing 2 to $\alpha_{c2} = 22°$ for wing 1.

Figure 2 plots the functions $c_y(\alpha)$ and $m_z(\alpha)$ for asymmetric-profile wing 1 at Re = 1.64×10^6, 2.02×10^6, and 2.5×10^6 when the angle of attack both increases and decreases. For these Reynolds numbers, the curves exhibit hysteresis: as Re grows, so do $c_{y\,max}$ and the critical angle of attack α_c (from $21°$ at Re = 1.64×10^6 to $24°$ at Re = 2.5×10^6), while the hysteresis loop area decreases. Similar behavior was observed previously for symmetric-profile wings [3–5].

For wing 1, the effect of Re on the polar $c_{ya} = f(c_{xa}, \alpha)$ and on the dependences $m_z(c_y)$ and $\bar{x}_C = f(\alpha)$ is demonstrated in Figs. 3 and 4, respectively. As before, the dependences are hysteretic. The loop area in the curve $m_z(c_y)$ decreases with increasing Re. The aerodynamic center position vs. angle-of-attack dependence $\bar{x}_C = f(\alpha)$ exhibits hysteresis in the range $\alpha = 18°–25°$. The relative position of the wing aerodynamic center ($\bar{x}_C = 0.26$) remains virtually unchanged up to $\alpha = 23°$ ($\dot{\alpha} > 0$), increases to $\bar{x}_C = 0.35$ at $\alpha = 22°–25°$, and then ($\alpha > 25°$) shifts back along the wing to $\bar{x}_C = 0.5$. As the angle of attack decreases to $\alpha = 20°$, the aerodynamic center returns to $\bar{x}_C = 0.35$ and again reaches the position $\bar{x}_C = 0.26$ in the range $\alpha = 18°–20°$.
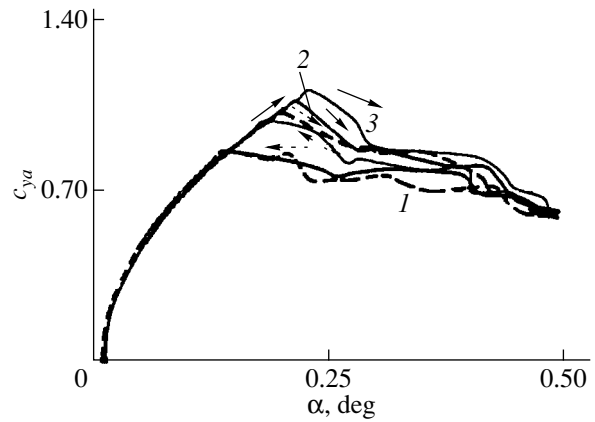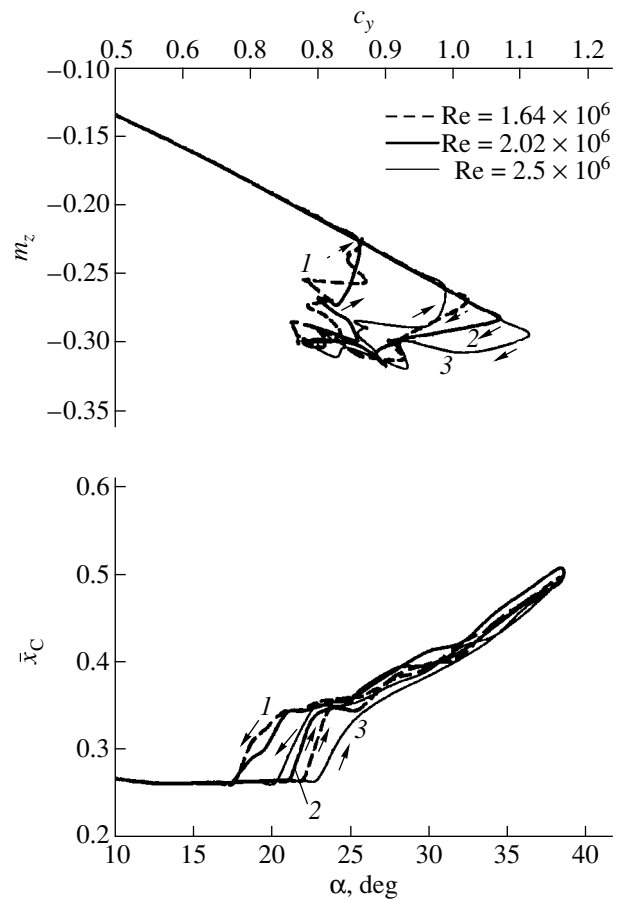

Fig. 4. Curves $m_z(c_y)$ and $\bar{x}_C = f(\alpha)$ for wing 1 with deflected and undeflected flaps (for *1–3*, see Fig. 2).
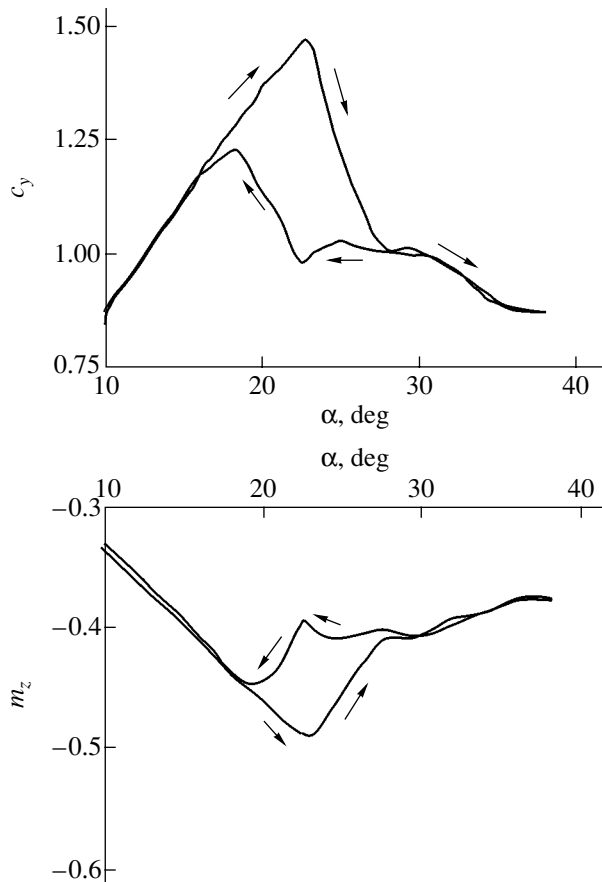
Fig. 5. Normal-force and longitudinal-moment coefficients vs. angle of attack for the asymmetric-profile wing with deflected flaps ($\delta = 20°$) at Re = 2.02×10^6.
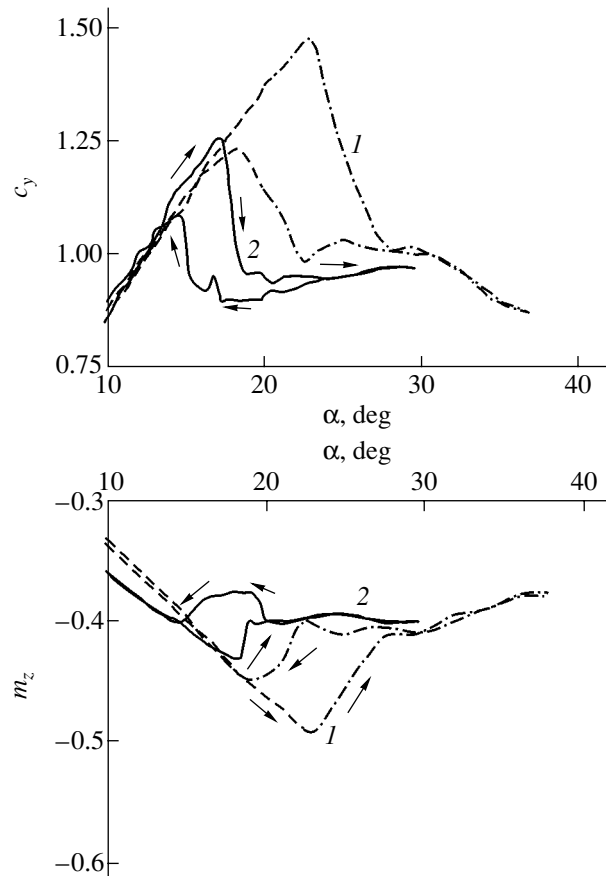

Fig. 6. Normal-force and longitudinal-moment coefficients vs. angle of attack for the (*1*) asymmetric- and (*2*) symmetric-profile wings with deflected flaps ($\delta = 20°$) at Re = 2.02×10^6.

Figure 5 shows the curves $c_y(\alpha)$ and $m_z(\alpha)$ obtained upon testing wing 1 with the flap deflected by $\delta = 20°$ (Re = 2.02×10^6). The deflection of the flap increases noticeably (roughly twofold) the hysteresis loop area compared with the area at $\delta = 0$ (Fig. 1). Because of the asymmetry of the profile, the flap deflection $\delta = 20°$ shifts the angle-of-attack range where hysteresis is observed to $\alpha = 17°–27°$ (cf. $\alpha = 15°–23°$ for the symmetric-profile wing, Fig. 6).

CONCLUSIONS

From our experimental dependences of the aerodynamic force coefficients and moment coefficients on the angle of attack for asymmetric- and symmetric-profile rectangular wings, one can draw the following conclusions. The profile curvature has a considerable effect on the size and shape of hysteresis loops in the curves of the static aerodynamic forces and moments. For the asymmetric-profile wing, the deflection of the flaps by $\delta = 20°$ increases the area of the loops in the curves $c_y(\alpha)$ and $m_z(\alpha)$ compared with these curves taken at $\delta = 0$.

REFERENCES

1. E. M. Jacobs, NACA Rep. 391 (1931).

2. A. E. Winkelmann and J. B. Barlow, AIAA J. **18**, 1006 (1980).

3. V. Ya. Neĭland and G. I. Stolyarov, Uch. Zap. Tsentr. Aérogidrodin. Inst. **13** (1) (1982).

4. Yu. A. Kolmakov, Yu. A. Ryzhov, G. I. Stolyarov, and V. G. Tabachnikov, Tr. Tsentr. Aérogidrodin. Inst., No. 2290 (1985).

5. V. G. Tabachnikov, Tr. Tsentr. Aérogidrodyn. Inst., No. 2290 (1985).

6. E. A. Karavaev, Yu. A. Prudnikov, and E. A. Chasovnikov, Uch. Zap. Tsentr. Aérogidrodin. Inst. **17** (6) (1986).

7. M. A. Golovkin, V. L. Gorban', E. V. Simuseva, and A. N. Stratonovich, Uch. Zap. Tsentr. Aérogidrodin. Inst. **18** (3) (1987).

8. I. V. Kolin, K. F. Latsoev, V. P. Mamrov, *et al.*, Preprint No. 87, TsAGI (Joukowski Central Institute of Aerohydrodynamics, 1996).

9. I. V. Kolin, V. P. Mamrov, T. I. Trifonova, and D. V. Shukhovtsov, in *Proceedings of the 5th Russian–Chinese Symposium on Aerodynamics and Flight Dynamics, 1997.*

10. I. V. Kolin, K. F. Latsoev, V. P. Mamrov, *et al.*, Preprint No. 114, TsAGI (Joukowski Central Institute of Aerohydrodynamics, 1998).

11. I. V. Kolin, V. L. Sukhanov, T. I. Trifonova, and D. V. Shukhovtsov, Uch. Zap. Tsentr. Aérogidrodin. Inst. **31** (1–2) (2000).

12. I. V. Kolin, V. K. Svyatodukh, T. I. Trifonova, and D. V. Shukhovtsov, Tr. Tsentr. Aérogidrodin. Inst., No. 2649 (2001).

13. I. V. Kolin, T. I. Trifonova, and D. V. Shukhovtsov, Zh. Tekh. Fiz. **71** (7), 128 (2001) [Tech. Phys. **46**, 914 (2001)].

14. I. V. Kolin, V. L. Sukhanov, T. I. Trifonova, and D. V. Shukhovtsov, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 2 (2002).

Translated by V. Isaakyan

SHORT
COMMUNICATIONS

Avalanche Breakdown Voltage of *n–p–n* Transistors in I²L Logic Elements

S. V. Shutov, A. N. Frolov, and A. A. Frolov

Institute of Semiconductor Physics, Kherson Division, National Academy of Sciences of Ukraine,
Kherson, 73008 Ukraine
Received April 16, 2003

Abstract—A method of determining the avalanche breakdown voltage of vertical switching *n–p–n* transistors in I²L logic elements is verified experimentally. © 2004 MAIK "Nauka/Interperiodica".

At the beginning of I²L logic circuits, punch-through breakdown (when the space charge region of the collector junction merges with the emitter through the base) was considered to be the basic breakdown mechanism in vertical switching-type *n–p–n* transistors. Accordingly, the punch-through voltage was taken to be equal to the collector–emitter breakdown voltage U_{ce} [1]. As the I²L technology matured and the area of I²L application expanded, designers arrived at the conclusion that not only the punch-through voltage but also the avalanche breakdown voltage should be determined. As is known [2], the breakdown voltage in drift and diffusion common-emitter bipolar transistors is given by

$$U_{ce} = \frac{U_{cb}}{\sqrt[n]{B_N + 1}}, \qquad (1)$$

where U_{cb} is the avalanche breakdown voltage of the collector–base (*p–n*) junction, B_N is the gain, and *n* is the exponent (*n* = 3 for *p–n–p* transistors and *n* = 4 for *n–p–n* transistors).

Our goal was to check the applicability of expression (1) for finding the avalanche breakdown voltage of vertical switching *n–p–n* transistors in I²L elements. For this purpose, we measured the breakdown voltage on 76(6KÉF0.4/380ÉKÉS0.01) single-layer epitaxial phosphine-doped silicon structures of different topology.

On each of the wafers, we fabricated single-collector *n–p–n* transistors with different rated currents and ratios S_c/S_e (where S_c and S_e are the surface areas of the collector and emitter, respectively) of 0.88, 0.75, and 0.63; double-collector *n–p–n* transistors with S_c/S_e = 0.24 + 0.24 and 0.40 + 0.06; and three- and four-collector *n–p–n* transistors with different and equal collector dimensions.

Four lots of four wafers each were fabricated and tested. The base region was formed by implantation of boron ions with various doses, followed by drive-in at 1150°C for different times *t*. The process conditions are given in the table. The collectors were formed by one-stage phosphorus diffusion at 1040°C. In reference areas of the wafers, the collector–base and emitter–base junction depths (X_{cb} and X_{eb}, respectively) were measured by the ring lap technique.

Subsequent to the opening of contact windows, vacuum evaporation of aluminum, and patterning of the aluminum layer by means of photolithography, the wafers were annealed in argon at 510°C for 25 min. In all the collectors of each of the elements at five sites on each of the wafers, we measured the gain B_N at a voltage of 0.5 V, the breakdown voltage U_{cb} of the collector–base junction, and the breakdown voltage U_{ce} with an L2-56 meter of semiconductor device parameters. Also, we determined the extrapolation voltage U_{ext} (Erli voltage) from the slope of the output *I–V* characteristics. The average values of these parameters are listed in the table.

It is seen that the extrapolation voltages U_{ext} for I²L elements of different topology that were fabricated on one wafer are the same. This is of no surprise, since the diffusion regions of the bases and collectors are formed simultaneously. Some of the *n–p–n* transistors in I²L elements have U_{ce} close to U_{ext}.

Since the punch-through voltage and the extrapolation voltage are given by the same expressions [3, 4], the voltage U_{ce} also equals the punch-through voltage if $U_{ce} = U_{ext}$. Then, lower values of U_{ce} are specified by avalanche breakdown. From (1), we find that

$$n = \log(B_N + 1)/\log(U_{cb}/U_{ce}). \qquad (2)$$

The values of *n* for the structures where U_{ce} is governed by avalanche breakdown are listed in the table. It is seen that *n* depends on the surface area ratio but is nearly independent of the process conditions.

From the dependence $n = f(S_c/S_e)$ obtained, one can derive an expression for *n* when expression (1) is used to calculate the avalanche breakdown voltage of I²L

SHUTOV *et al.*

Table

Lot no.	Process parameters		S_c/S_e	Measured electrical parameters				n calculated by (3)
	Q_b, μC/cm^2	t, min		B_N, u	U_{ce}, V	U_{cb}, V	U_e, V	
1	4	150	0.88	83	19.4	4.6	4.6	
			0.75	68	19.4	4.6	4.6	
			0.63	55	19.4	4.7	4.6	
			0.4	31	19.4	4.6	4.6	
			0.24	15	19.4	4.7	4.6	
			0.11	5.5	19.4	4.7	4.7	
2	12	105	0.88	42	12.2	5.3	8.7	4.51
			0.75	34	12.2	6.3	8.7	5.38
			0.63	27	12.2	7.2	8.7	6.32
			0.4	15	12.2	8.7	8.7	
			0.24	8	12.2	8.8	8.8	
			0.11	2.5	12.2	8.8	8.8	
3	40	60	0.88	40	8.3	3.6	7.8	4.45
			0.75	33	8.3	4.3	7.8	5.36
			0.63	27	8.3	4.9	7.8	6.32
			0.4	15	8.3	6.3	7.8	10.0
			0.24	7.5	8.3	7.3	7.9	16.6
			0.11	2.5	8.3	7.9	7.9	
4	100	40	0.88	35	6.9	3.1	7.2	4.48
			0.75	28	6.9	3.6	7.2	5.17
			0.63	22.5	6.9	4.2	7.2	6.38
			0.4	13	6.9	5.3	7.2	10.0
			0.24	6.5	6.9	6.1	7.3	16.6
			0.11	2.0	6.9	6.7	7.3	36.3

structures:

$$n = 4S_e/S_c. \qquad (3)$$

CONCLUSIONS

(i) The avalanche breakdown voltage of n–p–n transistors in I^{2}L elements depends not only on the avalanche breakdown voltage of the collector–base (p–n) junction and the transistor gain but also on the surface area ratio S_c/S_e.

(ii) Expression (1) can be universally used to calculate the avalanche breakdown voltage U_{ce}. However, while for n–p–n and p–n–p bipolar transistors the expo-

nent n is a constant, for I^{2}L structures, it depends on the surface area ratio (see (3)).

REFERENCES

1. N. A. Avaev, V. N. Dulin, and Yu. E. Naumov, *Large-Scale Integrated Injection Logic Circuits* (Sov. Radio, Moscow, 1977).
2. A. B. Greben, *Design of Analog Integrated Circuits* (Énergiya, Moscow, 1976).
3. I. P. Stepanenko, *Fundamentals of Microelectronics* (Sov. Radio, Moscow, 1980).
4. S. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981; Mir, Moscow, 1984), Vol. 1.

Translated by V. Isaakyan

Base Thickness Determination
from the Punch-Through Voltage for Vertical *n–p–n*
Transistors Incorporated into I²L Elements

S. V. Shutov, A. N. Frolov, V. N. Litvinenko, and A. A. Frolov

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine
Received April 16, 2003

Abstract—A method of determining the base thickness of *n–p–n* transistors in I²L elements from the punch-through voltage is verified on test I²L structures. © *2004 MAIK "Nauka/Interperiodica".*

In designing I²L circuits, as well as in quality control of the circuits during fabrication, there appears the need to determine the transistor base thickness. Specifically, the base thickness may be estimated from the punch-through voltage $U_{\text{p-th}}$, which is the voltage at which the space charge region (SCR) of the collector–base junction merges with the emitter.

The breakdown voltage U_{br} of common-emitter *n–p–n* transistors, as well as of any bipolar transistors incorporated into I²L circuits, equals the lower of two voltages values: the avalanche breakdown voltage U_{av} and the punch-through voltage $U_{\text{p-th}}$. The former is generally given by [1]

$$U_{\text{av}} = \frac{U_{\text{cb}}}{\sqrt[n]{B_N + 1}},\qquad(1)$$

where U_{cb} is the avalanche breakdown voltage of the collector–base *p–n* junction, B_N is the gain, and *n* is the structure-dependent exponent.

The punch-through voltage $U_{\text{p-th}}$ at which the SCR thickness becomes equal to the base thickness W_{b} (as determined by the metallurgical boundaries of the emitter and collector junctions) is found from the expression [2]

$$U_{\text{p-th}} = \frac{W_{\text{b}}^2 e N_{\text{b}}}{2\varepsilon\varepsilon_0} - \varphi_{\text{c}},\qquad(2)$$

where *e* is the electron charge, N_{b} is the impurity concentration in the base, ε is the permittivity of the semiconductor, ε_0 is the permittivity of free space, and φ_{c} is the contact potential difference at the collector junction.

If the base is doped uniformly, the determination of the punch-through voltage is not a matter of concern. The base thickness (or, more exactly, the minimal base thickness $W_{\text{b, min}}$ tailored for a desired breakdown voltage U_{br}) is also easy to find:

$$W_{\text{b, min}} = \sqrt{\frac{2\varepsilon\varepsilon_0(U_{\text{br}} + \varphi_{\text{c}})}{e N_{\text{b}}}}.\qquad(3)$$

However, in the standard I²L technology and in the combined bipolar–I²L technology, the base regions of *n–p–n* transistors are formed by ion doping or dopant predeposition with subsequent drive-in. In this case, the impurity distribution in the base takes an exponential form. As a result, use of expressions (2) and (3) poses certain difficulties, since it becomes necessary to know the average impurity concentration N_{b}^* in the base, which lies between the concentrations at the emitter–base and collector–base *p–n* junctions [3].

The standard I²L technology is based on drift bipolar *n–p–n* transistors. This technology is mature and quite compatible with the combined bipolar–I²L process. In the latter case, an *n–p–n* transistor of an I²L element may be viewed as an inversely operated *n–p–n* transistor, where its emitter acts as a collector and vice versa. Therefore, for an I²L element, the impurity concentration N_{eb} at the emitter–base *p–n* junction can be designated as N_{cbi} (the impurity concentration at the collector–base *p–n* junction upon inversion) and the concentration N_{cb}, as N_{ebi}.

The well-known technique of determining the punch-through voltage [4] for an inverted bipolar transistor implies the solution of the transcendental equation

$$\left(1 + \frac{\Delta_K}{L_{\text{a}}}\right)\exp\left(\frac{-\Delta_K}{L_{\text{a}}}\right) = 1 - \frac{(\varphi_{\text{c}} - U_{\text{p-th}})}{U_0\exp\left(\frac{\Delta_K}{L_{\text{a}}}\right)},\qquad(4)$$

where L_{a} is the characteristic diffusion length of accep-

Table 1. Experimental dimensions of diffusion regions, concentrations at the p–n junctions, and electrical parameters of n–p–n transistors incorporated into I²L elements

Wafer no.	Measurements					
	B_{Ni}	U_{bri}, V	U_{Ei}, V	X_{cbi}, µm	X_{ebi}, µm	W_b, µm
1	234	1.3	1.2	1.51	1.84	0.33
2	155	2.4	2.4	1.45	1.86	0.41
3	98	4.3	4.1	1.37	1.85	0.48
4	65	5.6	6.6	1.29	1.85	0.56
5	53	5.8	9.1	1.24	1.84	0.6
6	95	2.9	2.7	1.51	2.12	0.61
7	54	5.7	6.2	1.37	2.1	0.73
8	35	6.2	12.8	1.24	2.1	0.86

Table 2. Calculated results

Wafer no.	Calculated values		N_{bi}^* calculated by (7), cm^{-3}	W_b calculated by (3), µm
	N_{cbi}, cm^{-3}	N_{ebi}, cm^{-3}		
1	6×10^{16}	1.3×10^{16}	2.21×10^{16}	0.334
2	7.5×10^{16}	1.3×10^{16}	2.44×10^{16}	0.406
3	1×10^{17}	1.3×10^{16}	2.766×10^{16}	0.475
4	1.35×10^{17}	1.3×10^{16}	3.45×10^{16}	0.548
5	1.6×10^{17}	1.3×10^{16}	3.78×10^{16}	0.607
6	6×10^{16}	4.7×10^{16}	1.32×10^{16}	0.579
7	1×10^{17}	4.7×10^{15}	1.768×10^{16}	0712
8	1.6×10^{17}	4.7×10^{15}	2.356×10^{16}	0.863

tors in the base, Δ_K is the SCR width, and

$$U_0 = \frac{\varepsilon N_{ebi} L_a^2}{\varepsilon \varepsilon_0} \qquad (5)$$

is the design voltage [4].

The solution of the transcendental equation is much more complicated than direct calculation. Moreover, the solution of this equation requires that the base thickness, through which L_a is calculated, to be known *a priori*. Therefore, use of expression (4), which makes it possible to estimate the base thickness through the punch-through voltage, becomes a challenge. Theoretical considerations also show that the applicability of this expression is limited, especially when the impurity concentrations at the p–n junctions differ by no more than two to three times.

Expressions (2) and (3) are the most appropriate for forward and inverse calculations of the base thickness and punch-through voltage, since here only the concentration N_b^* must be known. For drift bipolar transistors,

this concentration is given by

$$N_b^*$$
$$= \exp\{\ln(N_{kb}N_{eb})/[2 - \ln(N_{eb}/N_{kb})/2\ln N_{eb}]\}. \qquad (6)$$

Since the punch-through voltage of a noninverted transistor is higher than that of an inverted device, the minus sign in the brackets should be changed to a plus. This change decreases the average concentration in the base and the associated punch-through voltage and gives a good fit to experimental data when the base is thin, i.e., when the impurity concentrations at the emitter and collector p–n junctions differ insignificantly (which is the case for I²L structures). However, as the base thickness and, accordingly, the concentration difference grow, the predicted values of the punch-through voltage and average concentration in the base tend to decrease, in conflict with the experimental evidence. Estimates for the SCR thickness [4] show that the associated expression must involve an additional factor when the SCR propagates from the high- to the low-concentration area (inverse operation) (compared with the case of opposite propagation). This factor will change the expression for the average concentration in the base for the inversely operated device (change of sign and the introduction of additional factors).

;In order to determine the average concentration in the base of an n–p–n transistor in an I²L element (i.e., of an inversely operated bipolar transistor), we studied the punch-through voltage versus design parameters of I²L n–p–n transistors. I²L structures were made on 76(6КЭF0.4/380ЭКЭS0.01) phosphine-doped single-layer silicon epitaxial structures (wafers 1–5) and on 76(7.5КЭF1.2/380ЭКЭS0.01) phosphine-doped single-layer silicon epitaxial structures (wafers 6–8) fabricated by the technique suggested in [5].

The base regions of all the wafers were formed by ion doping at a dose of 15 µC/cm², followed by high-temperature drive-in at 1150°C for 75 min. The collectors were formed by one-step diffusion of phosphorus at 1040°C. Such process conditions minimize the effect of collector diffusion on the impurity distribution in the base and affect the depth X_{ebi} of the emitter–base p–n junction insignificantly. The collector thickness X_{cbi} was varied by varying the phosphorus diffusion parameters. Phosphorus diffusion into wafers 6 and 1, 7 and 3, and 8 and 5 was accomplished simultaneously. Therefore, the impurity concentrations at the collector–base p–n junctions and the collector thicknesses in each of the pairs of wafers were the same. At reference sites of the wafers, we measured the depths X_{cbi} and X_{ebi} of the collector–base and emitter–base p–n junctions by the ring lap method and then calculated the base thickness W_b as the difference between X_{cbi} and X_{ebi}.

Once the base and collector regions had been formed and contact windows had been opened by photolithography, we measured the current gain B_{Ni} and the breakdown voltage U_{br} of the n–p–n transistors of an

I²L element, as well as determined the extrapolation (or Erli) voltage V_{Ei} from the projections of the output I–V characteristics onto the voltage axis. The measurements averaged over each of the wafers are listed in Table 1. From the process parameters known and the junction thicknesses determined, the impurity concentrations at the emitter–base and collector–base p–n junctions (N_{ebi} and N_{cbi}, respectively) were measured (see Table 1).

It follows from Table 1 that the values of U_{bri} and U_{Ei} are close to each other on high-gain wafers, while on wafers where the gain is low U_{bri} is lower than U_{Ei}. This is because the punch-through effect in the inverted device is observed in thin bases and a decrease in the gain implies that the base has thickened to the point where avalanche multiplication of carriers, rather than punch-through, becomes the basic mechanism of breakdown. Upon avalanche breakdown, direct measurement of the punch-through voltage is impossible. However, according to [6], the punch-through voltage and the Erli voltage are given by the same expression for both inverted and noninverted bipolar transistors. Therefore, the Erli voltage found experimentally may be used to estimate the base thickness. The processing of experimental data yields the following estimator for the average impurity concentration in the base of an inverted drift transistor:

$$N_{bi}^* = \exp\{\ln(N_{cbi}N_{ebi})$$

$$\times [\ln N_{cbi}/\ln N_{ebi}]^2/[(2 + 4.5\ln(N_{cbi}/N_{ebi})) \qquad (7)$$

$$\times \sqrt{\ln N_{ebi}/\ln N_{cbi}}/\ln N_{ebi}]\}.$$

With this expression, one can find the base thickness of an inverted transistor from the punch-through (or Erli) voltage or vice versa. For diffusion transistors,

where $N_{cbi} = N_{ebi} = N_b$, expression (7) yields $N_{bi}^* = N_b$. The calculated values of N_{bi}^* and the values of W_b found by (3) with N_{bi}^* estimated from (7) are listed in Table 2.

The values of W_b given in Tables 1 and 2 diverge by no more than several percent. This discrepancy is within the accuracy of the ring lap method, which was used to measure the diffusion depth. Thus, using expression (7) and the Erli voltage U_{Ei}, one may find the base thickness of both drift and diffusion inverted transistors from (3) irrespective of the breakdown mechanism. Since n–p–n transistors in I²L elements are, in essence, inverted bipolar transistors, this computational scheme may also be applied to I²L circuits.

REFERENCES

1. A. B. Greben, *Design of Analog Integrated Circuits* (Énergiya, Moscow, 1976).
2. A. F. Trutko, *Design of Transistors* (Énergiya, Moscow, 1971).
3. I. E. Maronchuk, A. N. Frolov, and S. V. Shutov, Izv. Vyssh. Uchebn. Zaved. Radioélektron., No. 11, 66 (2001).
4. N. A. Avaev, V. N. Dulin, and Yu. E. Naumov, *Large-Scale Integrated Injection Logic Circuits* (Sov. Radio, Moscow, 1977).
5. N. A. Samoĭlov, A. N. Frolov, and S. V. Shutov, Zh. Tekh. Fiz. **68** (10), 136 (1998) [Tech. Phys. **43**, 1262 (1998)].
6. S. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981; Mir, Moscow, 1984), Vol. 1.

Translated by V. Isaakyan

SHORT
COMMUNICATIONS

Emission Characteristics of a Lead Erosion Laser Plasma

A. K. Shuaibov, M. P. Chuchman, and L. L. Shimon

Uzhhorod National University, vul. Pidgirna 46, Uzhhorod, 88000 Ukraine
e-mail: ishev@univ.uzhgorod.ua
Received June 19, 2003

Abstract—The spectra and dynamics of the line emission of a lead erosion laser plasma at a distance of 1 mm from the target are investigated. The plasma is ignited in a vacuum (P = 3–12 Pa) with a pulse-periodic neodymium laser (τ = 20 ns, f = 12 Hz, $W = (1–2) \times 10^9$ W/cm^2, and λ = 1.06 μm). The data obtained are used to analyze the emission dynamics and the mechanism of formation of the laser plume. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The wide application of photoelectric converters based on lead compounds stimulates their investigation [1]. One of the methods for optimizing film photocells produced by laser sputtering is the spectroscopic diagnostics of the formation and expansion of laser plasma [2]. The physics of this phenomenon is even more important: although the hydrodynamics of the formation and expansion of a plume is still poorly understood, it plays an important role in a number of laser-assisted production processes [3]. To obtain more information on the parameters of an expanding erosion plasma, we performed a spectroscopic study of the emission from a laser-induced lead plume at a distance of 1 mm from the target surface in a vacuum of P = 3–12 Pa.

The laser plasma was produced by a pulse-periodic neodymium laser (τ = 20 ns, f = 12 Hz, $W = (1–2) \times 10^9$ W/cm^2, and λ = 1.06 μm). The emission from the laser plume in the spectral range of 200–600 nm was studied by two methods: recording of time-average spectra and investigation of the dynamics of the emission lines of the target material. The radiation intensity was measured accurate to 5%, and the time measurement error was 1–2 ns. The equipment and experimental technique are described in detail in [4, 5].

The table gives the relative intensities of the most important spectral transitions and their percentages in the plasma line spectrum. The radiation intensity is given with allowance for the recording system sensitivity (I/k_λ). The emission intensity was maximum in the PbII 220.4-nm, PbI 261.4-nm, PbI 280.2-nm, and PbI 405.7-nm lines, emitted from levels that are close to the bottleneck of the recombination flux of the atomic plasma component (E = 6.5 eV) and from the lower levels of lead atoms and ions. Their contribution to the spectrum is 41% of the total intensity. The close arrangement of the energy levels favors nonradiative cascade transitions of atoms downward upon three-

Intensity distributions and percentages in the spectrum of a lead erosion laser plasma

λ, nm	Atom, ion	Term	E_{up}, eV	I/k_λ, arb. units	$\Delta I/k_\lambda$, %
220.4	PbII	$7s^2S_{1/2}$	14.79	2.93	9.3
224.7	PbI	$p7d^3D_2$	6.48	0.55	1.7
239.4	PbI	$p7d^3F_3$	6.5	0.47	1.5
244.4	PbI	$p8s^3P_0$	6.04	0.36	1.2
244.6	PbI	$p8s^3P_1$	6.036	0.44	1.4
247.6	PbI	$p7s^3P_2$	5.97	0.48	1.5
257.7	PbI	$p7s^1P_1$	6.13	0.95	3.0
261.4	PbI	$p6d^3D_2$	5.71	3.07	9.7
266.3	PbI	$p7s^3P_2$	5.97	1.49	4.7
280.2	PbI	$p6d^3F_3$	5.74	3.92	12.4
283.3	PbI	$p7s^3P_1$	4.37	2.06	6.5
287.3	PbI	$p6d^3F_2$	5.63	1.57	5.0
326.1	PbII	$10s^3S_{1/2}$	21.29	0.50	1.6
357.3	PbI	$p7s^1P_1$	6.13	1.32	4.2
364	PbI	$p7s^3P_1$	4.37	1.51	4.8
368.3	PbI	$p7s^3P_0$	4.34	1.62	5.1
374	PbI	$p7s^3P_2$	5.97	1.18	3.7
401.9	PbI	$p6d^3F_3$	5.74	0.31	1.0
405.7	PbI	$p7s^3P_1$	4.37	3.03	9.6
427.5	PbII	–	18.89	0.68	2.2
438.7	PbII	–	18.89	0.50	1.6
478.8	PbII	–	–	0.07	0.2
500.5	PbI	$p7s^1P_1$	6.13	0.11	0.3
520.1	PbI	$p8s^3P_1$	6.04	0.25	0.8
530.7	PbII	–	21.55	0.16	0.5
536.7	PbII	–	18.88	0.27	0.9
560.9	PbII	–	17	0.95	3.0
571.4	PbII	–	21.39	0.18	0.6
576.8	PbII	–	21.34	0.29	0.9
298.7	PbII	–	20.79	0.39	1.2

body recombination to the level $E = 5.74$ eV ($p6d^3F_3$). Radiative transitions from this level produce the maximum contribution to the spectrum intensity, namely, 12.4% in the PbI 280.2-nm line.

For detailed analysis, Fig. 1 shows typical waveforms of the emission intensity of the lead atoms and ions from the laser plasma. The vertical line shows the instant corresponding to the end of the laser pulse.

The basic feature of the waveforms is a sharp increase and decrease in the intensity of ionic radiation (at $t = 70$ ns, the intensity is ~10% of the maximum level) and a slow increase and decrease in the intensity of atomic radiation. Most waveforms exhibit two pronounced maxima, which is also characteristic of other materials [6], although the intensities of some spectral atomic transitions slowly increase and reach their maximum at a time corresponding to the minimum intensity for most of the waveforms.

Such behavior of the intensity is difficult to explain. Taking into account that the zone in which the radiation is analyzed is close to the surface and that the melting temperature of the lead is rather low, this behavior can only be explained if we assume that the emission dynamics is mainly determined by the density of the evaporated material. This fact can explain the appearance of two maxima in most waveforms. The first emission maximum correlates with the time of laser action, and the second can be caused by the evaporation of heated inner layers of the material under the action of plasma radiation. Different positions of the maxima in different spectral lines can be related to the plasma expansion dynamics and the distributions of electron density and temperature, which affect the characteristics of spectral transitions whose spectroscopic lifetimes are greater than the resolution of our apparatus. Since the second maximum appears synchronously with the decrease in the transition wavelength, the contribution of nonradiative transitions is significant, but their role weakens with time.

If we consider this fact along with the emission dynamics of the PbI 357.3-nm line ($E_{up} = 6.13$ eV) with one broad maximum, then it should shift toward the maximum ionic emission intensity due to the recombination mechanism for population of the upper short-living level. The peak in the PbI 280.2-nm line emission, on the contrary, is clearly pronounced. It begins before the end of the laser pulse and corresponds to a combination of terms with a forbidden total orbital quantum number L: $6p^2\ ^3P_2$–$p6d^3F_3$. This transition can result from an external action, and its probability increases with increasing plasma density. Thus, we can conclude that electron–atom collisions are dominant in the plasma before the laser power decreases. The efficiency of this interaction can be traced by the intercombination emission of the PbI 520.1-nm line. The forbiddenness for the intercombination transition is less rigorous, and it is favored by free electrons. The first maximum of the radiation intensity closely resembles the shape of the
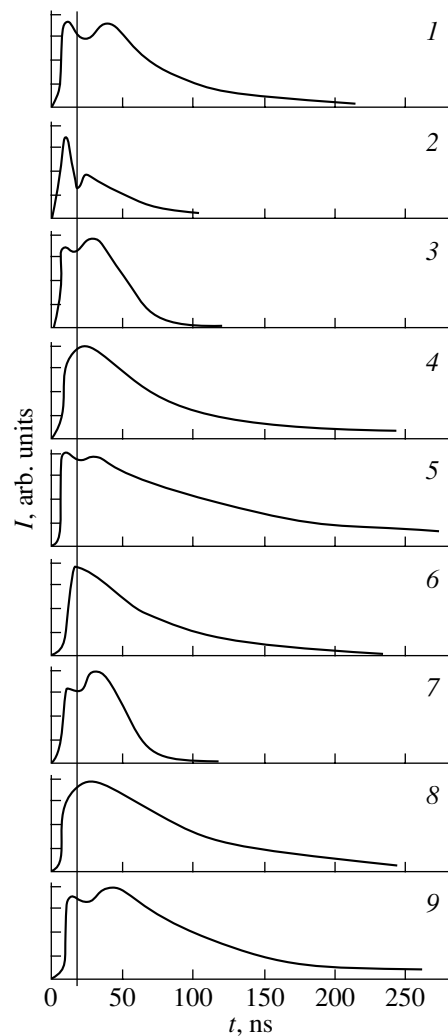

Fig. 1. Waveforms of the line emission intensity from a lead erosion laser plasma: (*1*) PbI 261.4 nm, (*2*) PbI 520.1 nm, (*3*) PbII 438.7 nm, (*4*) PbI 364.0 nm, (*5*) PbI 405.7 nm, (*6*) PbI 280.2 nm, (*7*) PbII 424.5 nm, (*8*) PbI 357.3 nm, and (*9*) PbI 368.3 nm.

laser pulse, and the second one appears on the leading edge of the second peak of allowed spectral transitions. This mechanism correlates with the data in [7].

Thus, the emission intensity increases with the probability of a spectral transition under given conditions and the probability of formation of particles in a certain energy state. Comparing the emission dynamics for transitions from the same upper level (Fig. 1, curves *4*, *5*), we can also conclude that heating is only due to radiation transfer. The transfer efficiency is substantially affected by nonradiative transitions. Since the lower states for both transitions are metastable, with $E_{low}(405.7$ nm$) > E_{low}(364.0$ nm$)$, the self-reversal for the PbI 405.7-nm transition can be more pronounced when atoms at the lower energy levels rise to a higher level, although the emission dynamics indicates that both transitions are forbidden in L.
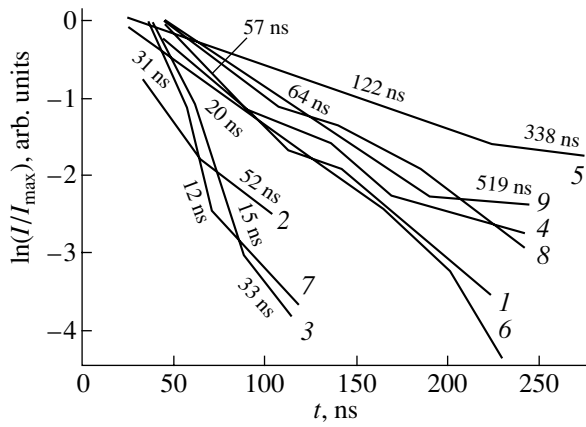
Fig. 2. Time dependence of the logarithm of the line emission intensity normalized to its maximum value (the curves are approximated by straight segments): (*1*) PbI 261.4 nm, (*2*) PbI 520.1 nm, (*3*) PbII 438.7 nm, (*4*) PbI 364.0 nm, (*5*) PbI 405.7 nm, (*6*) PbI 280.2 nm, (*7*) PbII 424.5 nm, (*8*) PbI 357.3 nm, and (*9*) PbI 368.3 nm.

To analyze processes in the late stages of expansion, we plot the time dependence of the logarithm of the intensity normalized to its maximum value (the curves are approximated by straight segments). The recombination times estimated from these curves are 52–57 ns for the Pb$^+$ ions and 12–15 ns for the Pb^{2+} ions in the time interval 50–100 ns. Until the 50th nanosecond, the dynamics of spectral transitions is strongly affected by the radiation transfer to the target surface. After $t = 100$ ns, the intensity variation again slows, which is clearly seen for the upper excited states of lead atoms. After $t = 200$ ns, the population of highly excited particles sharply decreases and the lower excited states form in a more steady manner. Probably, this is due to the association–dissociation excitation processes in a relatively cold and dense laser plasma.

Thus, in laser sputtering of lead-based films, ion-assisted processes and the specific features of energy redistribution in a laser plume are of great importance. These features are related to the multielectron valence atomic shell of lead and the related shape of the emission spectrum, as well as the evaporation dynamics and radiation transfer.

The emission dynamics is significantly affected by both the plasma expansion dynamics and the parameters of the expanding plasma. Their time dependences can be inferred from the waveforms of the spectral line intensities. According to this approach, the electron density conforms well to the shape of the laser pulse and increases again at the leading edge of the second evaporation peak. The hydrodynamic pressure is maximum at the end of the laser pulse and then gradually decreases. The recombination time of doubly charged lead ions is 12–15 ns, and that of singly charged ions is 52–57 ns.

The results obtained can be of interest for the optical spectroscopy of emission from a laser plasma of lead-based compounds and for the optimization of laser sputtering of such compounds.

REFERENCES

1. N. N. Syrbu and V. Z. Cebotari, J. Phys.: Condens. Matter **10**, 3467 (1998).
2. P. J. Wolf, J. Appl. Phys. **76**, 1480 (1994).
3. S. S. Harilal, C. V. Bindhu, M. S. Tillak, *et al.*, J. Phys. D **35**, 2935 (2002).
4. A. K. Shuaibov, L. L. Shimon, A. I. Dashchenko, *et al.*, Fiz. Plazmy **27**, 85 (2001) [Plasma Phys. Rep. **27**, 82 (2001)].
5. A. K. Shuaibov, L. L. Shimon, and M. P. Chuchman, Zh. Tekh. Fiz. **71** (5), 85 (2001) [Tech. Phys. **46**, 590 (2001)].
6. M. P. Chuchman, in *Proceedings of the Conference on Lasers, Applications and Technologies, Moscow, 2002*, p. 67.
7. Ya. N. Gnoevoǐ, A. I. Petrukhin, Yu. E. Pleshanov, and V. A. Sulyaev, Pis'ma Zh. Éksp. Teor. Fiz. **11**, 440 (1970) [JETP Lett. **11**, 297 (1970)].

Translated by K. Shakhlevich

SHORT
COMMUNICATIONS

Broadband Emission Spectrum of a Continuous Plasma Jet in a Mixture of Noble Gases with SF$_6$ Molecules

V. S. Rogulich and L. L. Shimon

Uzhhorod National University, vul. Pidgirna 46, Uzhhorod, 88000 Ukraine
e-mail: ishev@univ.uzhgorod.ua
Received March 4, 2003

Abstract—A broad band in the visible emission spectrum of a continuous Ar/Kr/SF$_6$ plasma jet is detected at a distance $l > 75$ mm from the plasmatron nozzle. Experimental dependences of the maximum emission intensity on the parameters of the plasma source are presented. © 2004 MAIK "Nauka/Interperiodica".

Earlier [1–3], we studied the optimum conditions for the formation of KrF*, XeF*, and ArF* excimer molecules in a continuous plasma jet ignited in a mixture of noble gases with SF$_6$ molecules and found broad visible emission bands, whose nature was not understood. This radiation was detected rather far ($l > 75$ mm) from the plasmatron nozzle.

In this work, we present the emission spectrum of an Ar/Kr/SF$_6$ mixture and basic experimental dependences of the maximum emission intensity of the band at $\lambda_{max} \approx 630$ nm on the parameters of the plasma source.

Experiments were performed with a plasmadynamic device in which the source of the plasma jet was a dc plasmatron with vortex stabilization of the electric arc and a sonic graphite nozzle with a critical diameter of ~5 mm operating at an arc current of 50–150 A and a discharge voltage of 15–35 V. Noble gas (He, Ar, Kr, or Xe) plasma was ignited in the plasmatron arc chamber, and SF$_6$ halide molecules were added to the noble gas plasma in both the prenozzle chamber and the formed plasma jet after the nozzle unit at various distances from it. The noble gas pressure in the arc and prenozzle chambers was varied from several dozens of pascals to 20 kPa. The radiation after the nozzle unit was recorded with an FÉU-106 photomultiplier and an MDR-2 monochromator. The equipment and experimental technique are described in detail in [1–4].

The emission spectra of the Ar/Kr/SF$_6$ jet was recorded in a broad wavelength range (200–800 nm). To avoid distortions of the spectra in the visible region due to the contribution of secondary radiation from the UV spectral region, in which KrF* and ArF* excimer molecules and OH radicals emit intensely, the spectral region with a wavelength $\lambda < 350$ nm was cut off with a light filter.

A typical emission spectrum of the Ar/Kr/SF$_6$ mixture consists of a broad emission band with an intensity maximum at $\lambda \approx 630$ nm and a half-width of ≈ 90 nm in the wavelength range 550–800 nm (Fig. 1). Several intense KrI lines exist against the background of the continuous spectrum. We studied the dependence of the maximum emission intensity of the band on the parameters of the plasma source. Figure 2 shows the dependence of the emission intensity of the band at $\lambda \approx 630$ nm on the flow rates of Kr and SF$_6$. In the absence of krypton or SF$_6$ molecules in the mixture, the emission intensity of the band is seen to be almost zero. The experimental dependences obtained indicate that the broadband emission exists only when krypton atoms and SF$_6$ gas molecules are simultaneously present in the mixture.

Figure 3 shows the dependence of the maximum emission intensity of the band in the Ar/Kr/SF$_6$ mixture on the distance along the axis of the plasma jet. The intensity is shown to be maximum at $l \approx 80$ mm from the plasmatron nozzle. As the plasmatron power increases, the radiation intensity of the band increases.

It should be noted that we also detected a broad emission band in the wavelength range 370–670 nm in a He/Xe/SF$_6$ mixture; however, experimental condi-
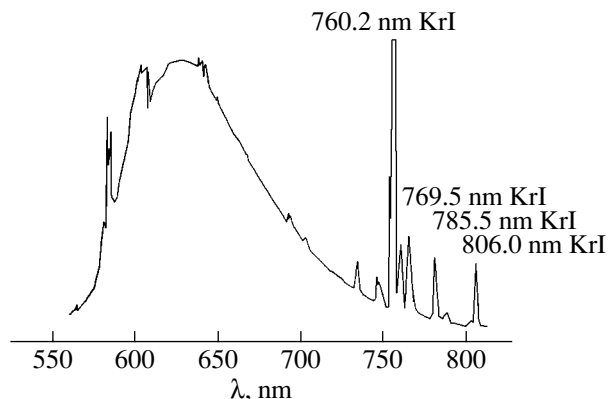

Fig. 1. Emission spectrum of the Ar/Kr/SF$_6$ plasma jet at $P \approx 8$ kPa, $W = 1.8$ kW, and $l = 81.5$ mm.

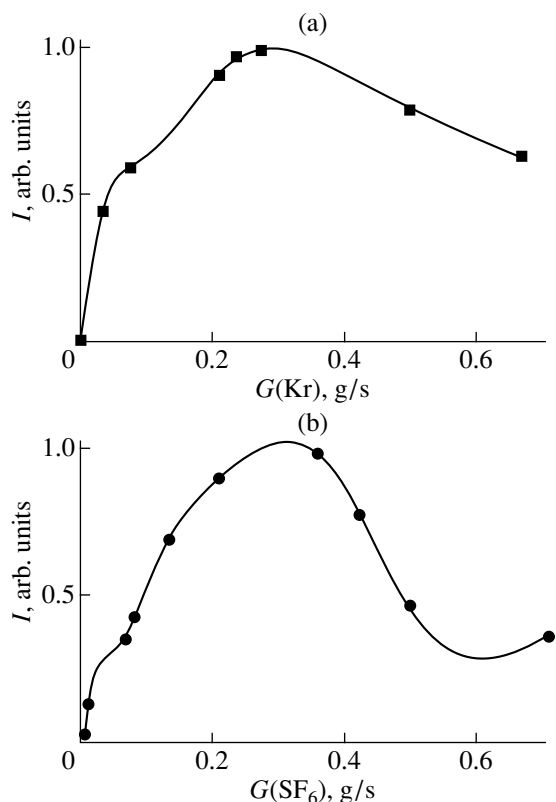**Fig. 2.** Emission intensity of the band at $\lambda = 630$ nm as a function of the partial flow rates of (a) krypton at $G(SF_6) = 0.24$ g/s and (b) SF_6 at $G(Kr) = 0.07$ g/s.

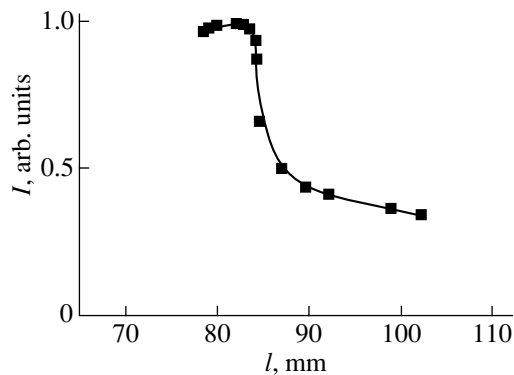**Fig. 3.** Variation in the maximum intensity of the band at $\lambda = 630$ nm with distance from the nozzle at $G(SF_6) = 0.1$ g/s, $G(Kr) = 0.24$ g/s, $P = 8$ kPa, and $W = 1.8$ kW.

tions were not optimized for this mixture, and the results require further examination and comprehensive analysis.

Based on the experimental data obtained and the analysis given above, we assume that an excited complex that consists of atoms or ions of the heavy noble gas and SF_6 molecules or their fragments and is more complex than the known di- and triatomic molecules efficiently forms in the $Ar/Kr/SF_6$ plasma jet rather far from the nozzle. It is not inconceivable that atoms of the plasma-forming gas (argon) are also involved in the formation of the complex. It should be noted that, under the conditions of high supercooling of the plasma beyond the nozzle (where the electron temperature is ~0.1 eV), the processes of electron attachment can occur in the plasma jet of the mixture of noble gases and strongly electronegative SF_6 molecules. They can produce negative SF_6^-, SF_5^-, F^-, and other fragments of the decomposition of the SF_6 molecule, which can serve as the formation sites of composite excited complexes (e.g., clusters).

As for other possible sources of broadband emission from the $Ar/Kr/SF_6$ mixture, we mention the triatomic excimer molecules ArKrF* and Kr_2F*, which are well-known in laser physics and have broad emission spectra. However, ArKrF* molecules emit in the UV spectral region, which was filtered out in the experiments, and the radiation of Kr_2F* molecules is mainly concentrated at $\lambda_{max} \approx 415$ nm. Therefore, these triatomic molecules are unlikely to be the main sources of the broadband emission from the $Ar/Kr/SF_6$ mixture.

REFERENCES

1. V. S. Rogulich, V. P. Starodub, and V. S. Shevera, Opt. Spektrosk. **69**, 756 (1990) [Opt. Spectrosc. **69**, 450 (1990)].
2. V. S. Rogulich and V. S. Shevera, Zh. Tekh. Fiz. **66** (4), 188 (1996) [Tech. Phys. **41**, 397 (1996)].
3. V. S. Rogulich and V. S. Shevera, Ukr. Fiz. Zh. **44**, 1082 (1999).
4. V. S. Rogulich and L. L. Shimon, Prikl. Fiz., No. 4, 60 (2002).

Translated by K. Shakhlevich

SHORT
COMMUNICATIONS

Improvement of the Thermoelectric Branch Efficiency
in the Case of the Linear Carrier Concentration Distribution

O. I. Markov

Orel State University, Orel, 302015 Russia
e-mail: Markov@e-mail.ru
Received July 3, 2003

Abstract—The one-dimensional boundary-value problem of determining the stationary temperature field of the thermoelectric branch is solved in the case of the maximal temperature difference. The temperature of the branch hot end is varied between 100 and 300 K. The calculation takes into account the Thomson effect, the distributed Peltier effect, and the temperature dependence of the charge carrier concentration. Optimization in terms of current and carrier concentration is performed. © 2004 MAIK "Nauka/Interperiodica".

Recent expansion of the production of thermoelectric coolers has emphasized the problem of increasing the thermoelectric figure of merit of semiconducting materials. This parameter is still far from the theoretical limit [1]. Use of thermoelectric branches of unequal length is one way of improving the thermoelectric figure of merit [2]. It was shown [3] that the figure of merit rises if the electrical conductivity increases and the thermal emf decreases from the hot to the cold end. Ivanova and Rivkin [4] used a more consistent approach based on the solution of the boundary-value problem of stationary heat conduction. Assuming that the thermopower, thermal conductivity, and electrical conductivity depend on temperature only slightly and the Thomson effect is negligible, they argued that a solution to this problem defines the optimal impurity distribution along the branch. The aim of this work is to solve the boundary-value problem [4] over a wide temperature range in the case of the linear carrier concentration distribution with allowance for the Thomson effect, the distributed Peltier effect, and the temperature dependence of the kinetic coefficients. The temperature dependence of the electrical conductivity is included through the temperature dependence of the mobility in the form $\text{const}\,T^{-3/2}$. A numerical solution to the problem is obtained in the case of maximal temperature difference.

Under the steady-state temperature difference with allowance for the Thomson effect, the temperature field of a one-dimensional adiabatically isolated thermoelectric branch with a nonuniform concentration distribution is given by the stationary heat conduction equation

$$\frac{d}{dx}\left(\chi\frac{dT}{dx}\right)+\frac{y^2}{\sigma}$$
$$-\frac{k}{e}yT\left(\frac{1}{n}\frac{dn}{dx}-\frac{3}{2T}\frac{dT}{dx}\right)=0 \tag{1}$$

with the boundary conditions

$$\chi\frac{dT}{dx}\bigg|_{x=0}=\alpha yT|_{x=0},\quad T|_{x=1}=T_1. \tag{2}$$

Unlike [4], we introduce here, for convenience, the quantity [5] $y = Jl/S$ as an optimization parameter.

For nondegenerate carrier statistics, the kinetic coefficients have the form

$$\sigma=enu,\quad u=\text{const}\,T^{-\frac{3}{2}},\quad \chi=\chi_1+2\left(\frac{k}{e}\right)^2 T\sigma,$$

$$\alpha=\frac{k}{e}\left(2+\ln\frac{2(2\pi mkT)^{3/2}}{nh^3}\right).$$

The lattice thermal conductivity χ_l, the proportionality factor const in the temperature dependence of the mobility, and the effective mass m were selected so that the thermoelectric figure of merit corresponded to semiconductors with $Z = 3.0 \times 10^{-3}$ K^{-1} at $T_1 = 300$ K. The carrier concentration along the branch varied, as in [4], by the linear law

$$n=n_0(1-gx). \tag{3}$$

The parameter $y = Jl/S$ may be called the reduced current [5].

Upon solving the boundary-value problem, the temperature difference was numerically optimized in terms of the reduced current and concentration n_0 at the cold end of the branch for a given g. The value of g was varied over wide limits: $0 \le g \le 0.999$, which corresponds to the variation of the ratio $k = n_0/n_1$ between the concentrations at the cold and hot ends in the range $1 \le k \le 10^3$.

The numerical solution to the problem stated by (1) and (2) is represented graphically. Figure 1 shows the maximal temperature difference versus hot end temper-
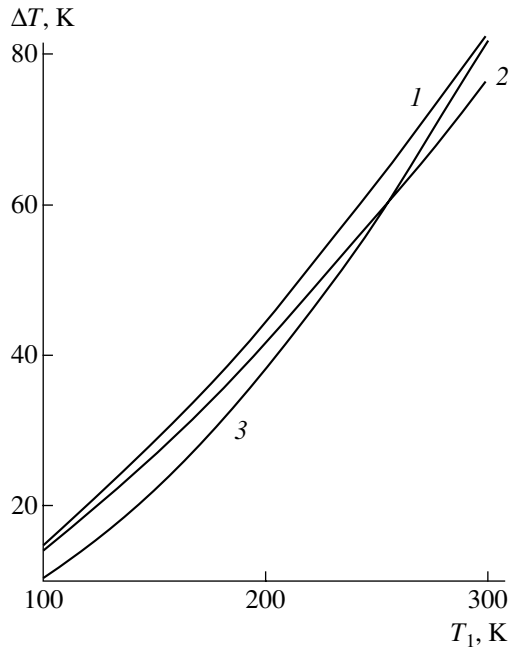
Fig. 1. Maximal temperature difference vs. hot end temperature: (*1*) solution to the problem stated by (1) and (2), (*2*) formula (4), and (*3*) solution to the boundary-value problem for bismuth telluride–based solid solutions.

ature under the conditions of maximal temperature difference. Curve *1* refers to the branch with the uniform concentration distribution ($g = 0$) and is obtained by
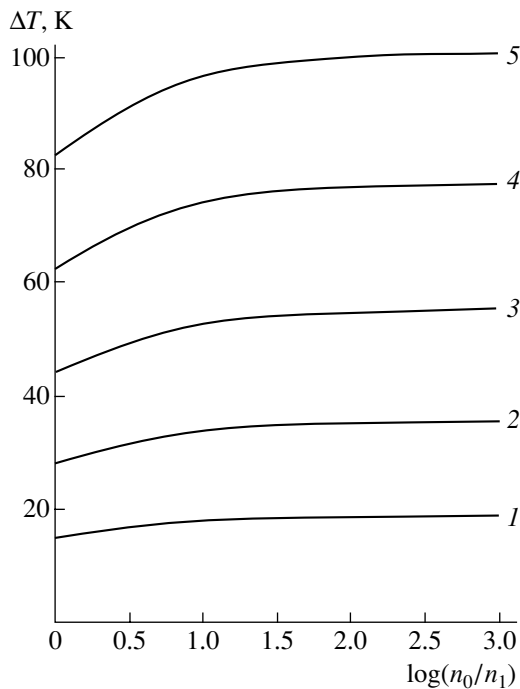
solving boundary-value problem (1), (2). Curve *2* reflects the temperature difference in the same branch that is calculated by the formula

$$\Delta T = T_1 + \frac{1}{z} - \sqrt{\frac{1}{z}\left(2T_1 + \frac{1}{z}\right)}. \qquad (4)$$

The discrepancy between these two curves is due to the Thomson effect and the temperature dependences of the kinetic coefficients. Curve *3* describes the temperature dependence of the maximal temperature difference for a bismuth telluride branch that has a thermoelectric figure of merit of 3×10^{-3} K^{-1}. In this case, the difference was calculated by solving the boundary-value problem [5] with the temperature dependences of the kinetic coefficients approximated by power polynomials. It is seen that the discrepancy between this curve and curves *1* and *2* grows as the hot end temperature drops, probably because of carrier degeneracy.

For branches with a concentration gradient, the maximal temperature difference increases. The maximal temperature difference between the cold and hot ends versus the logarithm of the concentration ratio k is shown in Fig. 2 for various temperatures of the hot end of the branch. From Fig. 2, it follows that it is hardly reasonable that the concentration ratio exceed eight to ten in the case of linear variation, since the temperature difference increases insignificantly when the concentration ratio is very high. For example, when the concentration ratio equals 10 (1000), the temperature dif-
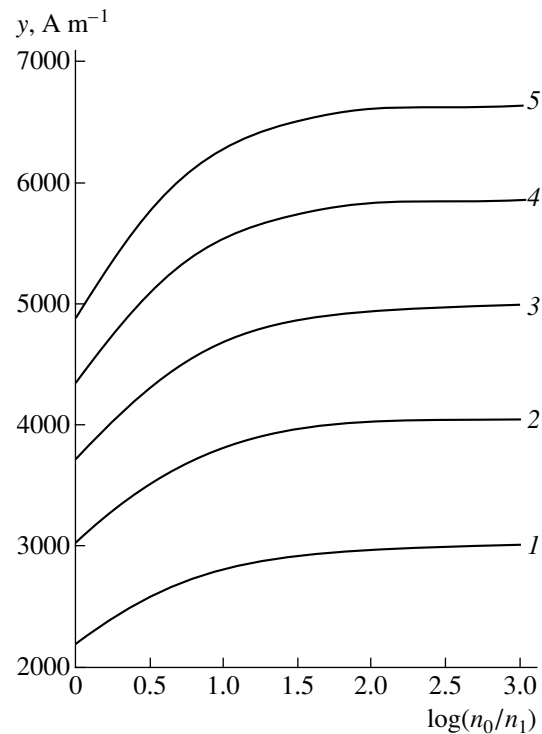

Fig. 2. Maximal temperature difference vs. the logarithm of the cold-to-hot end concentration ratio for a hot end temperature $T_1 =$ (*1*) 100, (*2*) 150, (*3*) 200, (*4*) 250, and (*5*) 300 K.

Fig. 3. Reduced current vs. the logarithm of the cold-to-hot end concentration ratio for various hot end temperatures. (*1–5*) Same as in Fig. 2.

ference increases by 16.8% (21.3%) for the hot end kept at 300 K. When the hot end temperature is 100 K, the temperature differences become 20.6 and 26.4%, respectively.

In the branch with the concentration gradient, the optimal current value also grows. Figure 3 shows the optimal value of the reduced current versus the logarithm of the ratio between the concentrations at the hot and cold ends. When the concentration ratio is 10 (1000), the current grows by 25.5% (36.7%).

The carrier concentration gradient decreases the cold end temperature because the Joule heat is compensated for (partially or completely) by the Thomson heat. Concentration optimization shifts the cold-end concentration n_0 toward higher values. Accordingly, the thermal emf at the cold end decreases, which causes the temperature difference to decline.

To conclude, classical statistics describes adequately the temperature dependence of the maximal temperature difference. Calculations show that graded-concentration thermoelectric branches may provide a much higher maximal temperature difference and improve the refrigerating capacity of thermoelectric coolers.

REFERENCES

1. E. S. Rittner, J. Appl. Phys. **30**, 702 (1959).
2. A. H. Boerijk, J. Appl. Phys. **32**, 1584 (1961).
3. V. A. Semenyuk, Teplofiz. Teplotekh., No. 35, 80 (1978).
4. K. F. Ivanova and F. S. Rivkin, Zh. Tekh. Fiz. **52**, 1406 (1982) [Sov. Phys. Tech. Phys. **27**, 853 (1982)].
5. O. I. Markov and O. A. Myznikova, in *Proceedings of the 8th International Seminar "Thermal Electronics and Applications," St. Petersburg, 2002* (Fiz.-Tekh. Inst. im. Ioffe, St. Petersburg, 2002).

Translated by V. Isaakyan

SHORT
COMMUNICATIONS

Uncooled Temperature Elements
Based on Semiconductor Crystals

I. M. Nesmelova, N. P. Cicina, and V. A. Andreev

Federal State Unitary Association NPO State Institute of Applied Optics,
Kazan, 420075 Tatarstan, Russia
e-mail: eugene@mi.ru
Received July 3, 2003

Abstract—The feasibility of fabricating InSb uncooled temperature-sensitive elements operating in the IR spectral range is considered. InSb electric parameters that are optimal for constructing bolometers with the highest possible sensitivity are evaluated. The results of computation are compared with the parameter values used in InSb temperature element prototypes. © 2004 MAIK "Nauka/Interperiodica".

In designing optoelectronic devices, the choice of a temperature-sensitive material that would provide a high device performance is one of the basic problems. At present, materials with high values of the temperature coefficient of resistance (TCR) and resistivity are the subject of extensive research, since they are viewed as candidates for uncooled high-sensitivity inexpensive bolometric photodetectors operating in the IR spectral range and compact precision quick-response temperature sensors. A variety of semiconductor materials with different values of the energy gap makes it possible to fabricate temperature detectors operating in interval 150–1000 K.

In the range of intrinsic conductivity, the TCR β of semiconductors,

$$\beta = -\frac{1}{R}\frac{dR}{dT} \quad \text{or} \quad \beta = \frac{1}{\sigma}\frac{d\sigma}{dT}, \tag{1}$$

is known to be negative, with its absolute value one order of magnitude higher than β of metals [1] (R is the resistance of the element and σ is the conductivity).

In this study, we demonstrate the feasibility of fabricating uncooled temperature detectors for optoelectronics with InSb single crystals. The basic electric parameters of InSb are calculated as functions of temperature and charge carrier concentration at near-room temperatures. The aim is to find the material parameters that are the most appropriate for the production of temperature (bolometric) elements with maximum possible sensitivity.

In the ranges of intrinsic and mixed conductivity, the electron and hole components of the conductivity add up,

$$\sigma = e(n\mu_n + p\mu_p), \tag{2}$$

and the Hall constant is given by

$$|R_{\mathrm{H}}| = \frac{1}{e}\frac{n\mu_n^2 - p\mu_p^2}{(n\mu_n + p\mu_p)^2}, \tag{3}$$

where n, μ_n and p, μ_p are the concentrations and mobilities of electrons and holes, respectively.

Let us introduce a new parameter $p = an_i$, where n_i is the intrinsic charge carrier concentration at a fixed temperature and a is a factor characterizing the electron–hole ratio in the material. Then, Eqs. (2) and (3) take the form

$$\sigma = en_i\mu_p(b/a + a), \tag{4}$$

$$|R_H| = \frac{1}{en_i}\frac{b^2 - a^2}{(b + a^2)^2}, \tag{5}$$
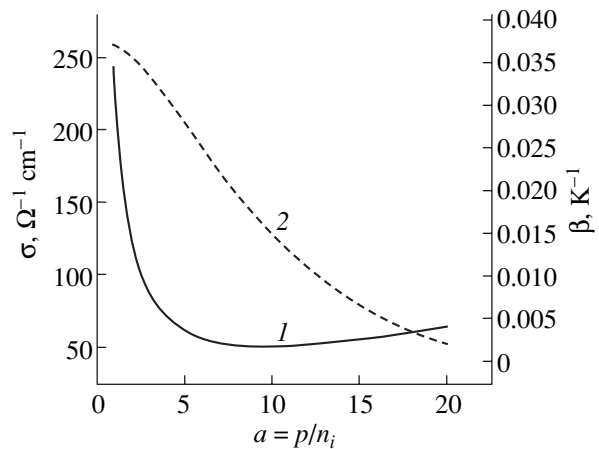
where $b = \mu_n/\mu_p$.

Fig. 1. (*1*) Electrical conductivity and (*2*) TCR absolute value vs. the parameter $a = p/n_i$ for InSb at 300 K.
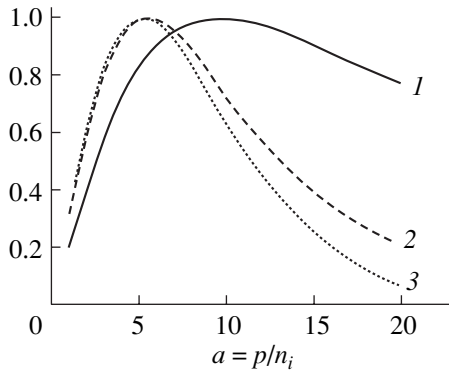
Fig. 2. (*1*) Sensitivity *S* of the bolometer, (*2*) Hall constant $|R_H|$ of single-crystal InSb, and (*3*) resistivity ρ vs. the parameter $a = p/n_i$ at 300 K.

A basic parameter of a temperature element is the voltage sensitivity *S*. For bolometers, $S \sim \beta R \alpha I$ [2], where α is the absorption coefficient of the sensitive element material and *I* is the current passing through the bolometer. For semiconductors in the range of intrinsic absorption, $\alpha = 10^3 – 10^4$ cm^{-1}. The voltage sensitivity of a bolometer may be approximated by the expression

$$S \approx \beta R = \frac{1}{\sigma^2}\frac{d\sigma}{dT}. \tag{6}$$

The intrinsic carrier concentration in InSb was calculated by the formula [3]

$$n_i = 2.9 \times 10^{11}(2400 - T)^{0.75}(1 + 2.7 \times 10^{-4}T)T^{1.5}$$
$$\times \exp[(-0.129 + 1.5 \times 10^{-4}T)/kT] \text{ (cm}^{-3}). \tag{7}$$

Electric parameters of InSb samples intended for uncooled temperature-sensitive elements (300 K)

Sample no.	$|R_H|$, cm^3 C^{-1}	σ, Ω^{-1} cm^{-1}	β*, K^{-1}	S_{rel}^{**}
1	903	33.3	-2.80×10^{-2}	1.00
2	725	34.0	-2.54×10^{-2}	0.90
3	680	29.6	-2.33×10^{-2}	0.94
4	332	25.6	-1.30×10^{-2}	0.61
5	140	24.0	-0.90×10^{-2}	0.45

* TCR was measured in the temperature range 307–315 K.
** S_{rel} are normalized to S_{max} of sample 1.

The carrier mobilities were found by the formulas [4]

$$\mu_n = 77\,000(T/300)^{-1.66} \text{ [cm}^2/(\text{V s})], \tag{8}$$

$$\mu_p = 2.55 \times 10^7 T^{-1.81} \text{ [cm}^2/(\text{V s})]. \tag{9}$$

Figure 1 shows the analytical dependences of the conductivity and TCR on the parameter $a = p/n_i$ for InSb at 300 K. The minimal value of σ, σ_{min}, is observed at $a = 9.59$ and lies in the range of mixed conductivity where $p > n$. The room-temperature intrinsic concentration in InSb, which was calculated by formula (7), is $n_i = 1.95 \times 10^{16}$ cm^{-3}. It follows that the material with $p = 1.87 \times 10^{17}$ cm^{-3} has the minimal electrical conductivity at 300 K.

Figure 2 shows S, $|R_H|$, and the resistivity $\rho = 1/\sigma$ at 300 K as functions of the parameter *a*. Each of the quantities is normalized to its maximum value. The maximal sensitivity S_{max} of the bolometer is observed at $a = 5.17$ ($p_{300\,K} = 1.01 \times 10^{17}$ cm^{-3}). The maximum value of $|R_H|$ corresponds to $a = 5.51$ ($p_{300\,K} = 1.07 \times 10^{17}$ cm^{-3}).

Thus, the maximal sensitivity of uncooled temperature elements made of InSb may be expected in acceptor-type samples. At 300 K, these samples must have the following parameters: $p = 1.01 \times 10^{17}$ cm^{-3}, $\sigma = 60.1$ Ω^{-1} cm^{-1}, $\rho = 1.7 \times 10^{-2}$ Ω cm, and $|R_H| = 989$ cm^3 C^{-1}. In such samples, $\beta = -2.77 \times 10^{-2}$ K^{-1}.

Similarly, we calculated the parameter *a* corresponding to the maximum of *S*, S_{max}, at a fixed temperature in the range 290–320 K. The position of S_{max} shifts insignificantly with temperature: from $a = 5.170$ (at $T = 290$ K) to 5.179 (at $T = 320$ K). The hole concentration rises from 8×10^{16} to 1.6×10^{17} cm^{-3}. The values of S_{max} and $|R_H|_{max}$ decrease by a factor of 2, and the TCR declines from $|\beta| = 2.95 \times 10^{-2}$ to 2.45×10^{-2}.

InSb single crystals were used to fabricate compact quick-response precision temperature elements (detectors) operating in the range 307–315 K (the range of a clinical thermometer). The temperature elements were made in the form of a 50-μm-thick meander and had an area of 1.5×1.5 mm^2. The resistance of the element was equal to 2.0–2.5 kΩ. The electrical parameters of the InSb samples from which the temperature sensors were fabricated are listed in the table. As is seen, the experimental values of the conductivity are lower than the calculated values, possibly because the mechanisms responsible for charge carrier scattering were not all taken into account in the calculations.

As follows from the tabulated data, bolometers made of InSb crystals with parameters close to the prescribed values had a sensitivity and TCR that are also close to the predicted values (samples 1–3). The param-

eters of samples 4 and 5 differ from the calculated results. Therefore, the elements made of these crystals had low values of S_{max} and TCR.

Thus, we calculated the optimal electrical parameters of InSb samples for fabricating uncooled temperature elements (bolometers and temperature detectors) for optoelectronics. It is shown that the maximum sensitivity of the elements can be obtained on acceptor-type crystals with $p_{300\,K} \sim 10^{17}$ cm^{-3}. Similar calculations may be performed for semiconductors with other values of the energy gap in which intrinsic conductivity falls into another temperature range. In this case, temperature elements can be designed for another spectral range.

REFERENCES

1. F. Lineveg, *Engineering Temperature Measurements* (Metallurgiya, Moscow, 1980).
2. V. Yu. Zerov and V. G. Malyarov, Opt. Zh. **68**, 88 (2001).
3. M. Oszwaldowski and M. Zimpel, J. Phys. Chem. Solids **49**, 1179 (1988).
4. O. Madelung, *Physics of III–V Compounds* (Wiley, New York, 1964; Mir, Moscow, 1967).

Translated by Yu. Vishnyakov