Acoustical Physics, Vol. 51, No. 1, 2005, pp. 1–4. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 5–8.

The *Akusticheskii Zhurnal* (*Acoustical Physics*) Is 50 Years Old

O. V. Rudenko
(Editor-in-Chief)

The first issue of the *Akusticheskii Zhurnal* appeared in 1955. Almost simultaneously, its English version was started under the title of *Soviet Physics–Acoustics*. The journal has become popular among acousticians of many countries all over the world. Its popularity was based on the high scientific standard that was set by the very first issues. The authors whose papers were published in the first volume included academicians N.N. Andreev and L.M. Brekhovskikh; the winner of the Nobel Prize Academician V.L. Ginzburg; professors G.D. Malyuzhinets, L.D. Rozenberg, and L.A. Chernov; and many other prominent scientists.

The first editorial board was approved by the Presidium of the Academy of Sciences of the USSR and included N.N. Andreev (editor-in-chief), L.M. Brekhovskikh, V.S. Grigor'ev, S.N. Rzhevkin, L.D. Rozenberg, and S.Ya. Sokolov. The merits of these outstanding scientists and their roles in the formation of the journal as a leading Soviet and, later, Russian periodic edition in acoustics were described in detail in the paper written by L.M. Lyamshev on the previous (40th) anniversary of the *Akusticheskii Zhurnal* (*Acoustical Physics*) (Akust. Zh. **41** (5), 677 (1995); Acoust. Phys. **41** (5), 595 (1995)).

The editorial board was periodically changed; new members that came to work in the journal were widely known scientists working in physical and applied acoustics. Beginning in 1962, the editor-in-chief was V.S. Grigor'ev, and since 1984, L.M. Lyamshev.

After the USSR was broken into independent states, Russian scientific journals and the book-publishing business fell into a difficult situation. The situation improved when the MAIK "Nauka/Interperiodica" publishing company started the commercial English version of the journal under the title *Acoustical Physics*; the corresponding decree of the Presidium of the Russian Academy of Sciences was pronounced in 1994. Since that time, the English version has been distributed outside Russia by the American Institute of Physics. It supports the Russian original version, which is the common product of the "Nauka" Academic Publishing Center and the MAIK "Nauka/Interperiodica" Publishing. The operation of these two companies is described in detail on the internet (http:/www.maik.ru).

The MAIK "Nauka/Interperiodica" regularly conducts competitions to determine the best publications of the year (papers or cycles of papers); the prizes are awarded by a special commission.

The MAIK prize of 1995 was received by Yu.M. Sukharevskii for his paper "Statistics of Basic Acoustic Parameters of Deep-Water Oceanic Regions and the Probabilistic Range of Sonar Systems" (Akust. Zh. **41** (5), 848 (1995); Acoust. Phys. **41** (5), 749 (1995)).

In 1996, the prize was given to Yu.I. Bobrovnitskii and T.M. Tomilina for the paper "General Properties and Fundamental Errors of the Method of Equivalent Sources" (Akust. Zh. **41** (5), 737 (1995); Acoust. Phys. **41** (5), 649 (1995)).

V.G. Andreev, V.N. Dmitriev, Yu.A. Pishchal'nikov, O.V. Rudenko, O.A. Sapozhnikov, and A.P. Sarvazyan received the prize of 1997 for the paper "Observation of Shear Waves Excited by Focused Ultrasound in a Rubber-like Medium" (Akust. Zh. **43** (2), 149 (1997); Acoust. Phys. **43** (2), 123 (1997)).

In 1998, the prize was given to L.M. Lyamshev and Yu.P. Lysanov for the paper "Sound Scattering by Random Volume Inhomogeneities with a Fractal Spectrum" (Akust. Zh. **44** (4), 506 (1998); Acoust. Phys. **44** (4), 434 (1998)).

An international team including A.P. Brysev, F.V. Bunkin, L.M. Krutyanskii, V.L. Preobrazhenskii, A.D. Stakhovskii, Yu.V. Pyl'nov, L. Hamilton, K. Cunningham, and S. Younghouse received the prize of 1999 for the paper "Nonlinear Propagation of a Quasi-Plane Conjugate Ultrasonic Beam" (Akust. Zh. **44** (6), 738 (1998); Acoust. Phys. **44** (6), 641 (1998)).

V.A. Zverev received the prize of 2000 for the paper "Acoustic Dark Field" (Akust. Zh. **46** (1), 75 (2000); Acoust. Phys. **46** (1), 62 (2000)).

In 2001, the prize was given to S.V. Krivokhizha, I.L. Fabelinskii, and L.L. Chaikov for the paper "Particular Features of the Acoustic Wave Propagation in Condensed Media in Different Conditions" (Akust. Zh. **47** (2), 238 (2001); Acoust. Phys. **47** (2), 194 (2001)).

The prize of 2002 was given to E.A. Kopyl, Yu.P. Lysanov, and L.M. Lyamshev for the paper "Sound Scattering by Random Fractal Inhomogeneities in the Ocean" (Akust. Zh. **48** (4), 517 (2002; Acoust. Phys. **48** (4), 453 (2002)).

The MAIK prize is not only a financial support for the winners but also a stimulus to an improvement of the quality of publications for other authors. Precisely

1063-7710/05/5101-0001$26.00

the quality of publications determines the demand for the journal among the international scientific community and eventually provides the necessary amount of subscriptions to support the publication of the Russian and English versions.

To evaluate the scientific quality of the *Akusticheskii Zhurnal* (*Acoustical Physics*) compared to other Russian journals devoted to research in physics, let us consider some objective estimates.

An important bibliometric factor characterizing the use of the journal by the international scientific community is the impact factor (Ip). It is calculated as follows: the total number of citations of the journal that occur in the literature within the current year and refer to the papers published within the two preceding years should be divided by the total number of papers published in the journal within these two years. Thus, Ip = 1 means that the papers published in the two preceding years were cited on the average once apiece in the current year.

The impact factor is calculated using the database published in the *Journal Citation Reports* by the Institute for Scientific Information, USA. In 2000, the database covered approximately 5700 journals, and only about 100 of them were Russian (for all natural sciences). The physical journals with the maximum Ip values were the *Journal of Experimental and Theoretical Physics* (Ip = 1.187), the *JETP Letters* (Ip = 1.411), and the *Physics–Uspekhi* (Ip = 1.182). The *Akusticheskii Zhurnal* (*Acoustical Physics*) occupied an intermediate place with Ip = 0.356; in 2001, this value was Ip = 0.363, and in 2002, Ip = 0.447. Thus, the Ip of the Russian physical journals is fairly stable. An exception is the *Physics–Uspekhi*: in 2003, the impact factor of this journal increased to Ip = 2.595, and, according to this value, the *Physics–Uspekhi* has become one of the ten most cited journals in the world.

However, the comparison of the *Akusticheskii Zhurnal* (*Acoustical Physics*) with the leading Russian physical journals is incorrect for two reasons.

First, in acoustics, many of the papers are theoretical ones of a "classical" type or of a methodical character. Citations of such papers are usually delayed and occur on the average within more than two years after their publication.

Second, acoustics is closely related to applied physics where the scientific priority problems are not as important as in basic physics. If a paper is devoted to a new physical effect or to a discovery of some unknown properties of matter or unknown processes in the universe, it is of global significance and, as a rule, is much cited. The authors of high-level (close to Nobel-Prize quality) basic works usually do not estimate their scientific achievements in terms of finances. By contrast, in applications, the priority is certified by an inventor's certificate and the success is estimated by the financial criterion. Ideas that are most valuable for practical applications are either not published in the journals or published with large delays. If an inventor publishes an innovative result too early, the idea may be put to practice by other people; in this case, the author evidently cannot expect even formal references to his paper.

To compare the journals devoted to different fields of science with allowance for their specificity, the so-called standard impact factor **K** was developed (see the paper "Bibliometric Evaluation of Russian Natural-Science Journals" by I.V. Marshakova-Shaikevich: Vestn. Ross. Akad. Nauk **73** (9), 788 (2003); Herald of the Russian Academy of Sciences **73**, No. 9, (2003)). The factor **K** is calculated as follows. First, a list of five journals representing a given area of research (acoustics in our case) and characterized by the highest individual values of the impact factor Ip is made up. Then, for the whole set of papers published in these five best journals, the characteristic factor for the given area of research, Ig, is calculated by the method of calculating the Ip factor. For acoustics in 2000, Ig = 1.52. The standard impact factor for a given journal is defined as **K** = (Ip/Ig) × 100%. According to this criterion, the *Akusticheskii Zhurnal* (*Acoustical Physics*) with **K** = 23.40 surpasses the *Journal of Experimental and Theoretical Physics* (**K** = 17.85), the *JETP Letters* (**K** = 21.22), and the *Physics–Uspekhi* (**K** = 17.64). According to the dynamics of the factors Ip and **K**, the *Akusticheskii Zhurnal* (*Acoustical Physics*) surpasses all other journals: from 1982 to 2000, its Ip increased from 0.094 to 0.356, and **K**, from 7.7 to 23.40.

Thus, our authors should not complain that their works are seldom cited abroad. The journal is being read, the scientific results and their authors are known to specialists in acoustics, and the influence of the *Akusticheskii Zhurnal* (*Acoustical Physics*) on the development of acoustics in the world is no smaller than that of any other Russian physical journal on the corresponding area of research.

The current editorial board and the editorial council (see the lists on the cover of the journal) were approved by the Bureau of the Division of Physical Sciences of the Russian Academy of Sciences on March 5, 2003. By the initiative of academicians heading the institutes that are involved in acoustical research, the editorial board was considerably renewed. The academicians themselves became members of the editorial council to determine the strategy of the journal. As a result, one can say that the current editorial council mainly consists of "stars" while the editorial board mainly consists of "workers." The regular monthly meetings of the editorial board are always attended by almost all its members. A reasonable ratio of the number of more experienced members to that of younger ones contributes to the successful (as we believe) work of the editorial board.

Now, I would like to present some new tendencies in the work of our editorial board to our authors and readers. Our aim is to raise the interest in the journal and to improve the scientific contents of the papers published

in it. Our actions are based on the ideas put forward by our Russian and foreign colleagues. In particular, from our colleagues working in the USA, we regularly receive evaluations of the contents of papers, the technical appearance, and the quality of translations for individual issues of the journal. The system of evaluation is organized by the American Institute of Physics (AIP).

Along with the regular issues of the journal, the editorial board decided to form specialized issues devoted to topical problems. Examples are the two first issues of 2003, which are devoted to the memory of L.M. Lyamshev, and issue 3 of 2004, which contains a set of papers on biological acoustics (in honor of the birthday of N.A. Dubrovsky). Currently, we are preparing an issue devoted to the scientific activities of V.A. Zverev and an extra issue devoted to the topical problems of seismoacoustics. The present issue, which is devoted to the 50th anniversary of the journal, is also an example of this kind of journal. For participation in specialized issues we invite well-known authors, and this strategy allows us to raise the average scientific quality of the papers. Specialized issues may be of interest to readers, first, because they are devoted to topical problems and, second, because the authors of papers concerned with one area of research and published in one issue may compete with each other in the significance of their results. It should be noted that our referees from the AIP acknowledged the high quality of the specialized issues that appeared in 2003.

We are trying to extend both the scope of our journal and the circle of well-known and actively working scientists willing to submit their papers for publication in it. Historically, such areas of research as underwater acoustics, nonlinear acoustics, and the theory of propagation and diffraction of acoustic waves predominated in the journal. At the same time, other fields of acoustics, such as the acoustics of hearing, speech acoustics, musical and architectural acoustics, bioacoustics, etc. were insufficiently represented. However, this does not mean that the editorial board will accept papers concerned with these topics but of a poor scientific quality. We intend to ask well-known specialists to write some papers and reviews for our journal.

Areas of research that deserve special consideration are those lying at the "boundaries" between acoustics and other physical and natural sciences, such as geophysics, mechanics, condensed-matter physics, materials science, and mathematical physics and simulations. Most publications concerned with these areas of research appear in other specialized journals, which have their own circles of readers. Thus, the exchange of information is often ineffective. Sometimes, this leads to doubling of certain studies or even of directions of research. Our editorial board intends to stimulate the mutual influence of the journals by cooperating with some new reviewers and authors specializing in the "boundary" areas of research and initially participating in other journals. Again, the novelty of the subjects must be accompanied by a high scientific quality and an appropriate style of publication to provide for an easy understanding of the papers by our readers. In our experience, in some cases, relatively poor papers rejected by other journals were submitted to the *Akusticheskii Zhurnal* (*Acoustical Physics*) under the pretense that their subjects were new to acoustics. We welcome unconventional papers but only if they are of a sufficiently high quality.

We also ask the new authors to cite as widely as possible the papers from the *Akusticheskii Zhurnal* (*Acoustical Physics*), as well as from JASA and other leading journals in acoustics. Even if formal citations seem to be unnecessary, i.e., not directly related to the results, these references may help the readers to establish the analogies with the neighboring fields of acoustics. It would be natural to assume that, if a paper contains no references to the papers published in the *Akusticheskii Zhurnal* (*Acoustical Physics*), it may be of no interest to our readers.

Despite the efforts of the editorial board and the professional skills of our reviewers, sometimes papers of questionable novelty or with a questionable reliability of the results appear in the journal. To raise the responsibility of the authors, we intend to welcome critical comments on our publications. In the section called "DISCUSSION", we intend to publish the critical comments and the responses of the authors (if any) to this criticism.

At the end of 2003, the Presidium of the Russian Academy of Sciences pronounced a Decree concerning Scientific Journals, which said: "The founder of a journal of the Russian Academy of Sciences (RAS) is the Russian Academy of Sciences. A journal of the RAS operates under the supervision of the Presidium of the RAS. The editor of a journal of the RAS is the Russian Academy of Sciences represented by the "Nauka" Academic Publishing Center of the RAS." The *Akusticheskii Zhurnal* (*Acoustical Physics*) is a journal of the Division of Physical Sciences, which, once every five years, presents a candidate for its editor-in-chief to the Presidium of the RAS for approval and approves the members of the editorial board. The editorial board makes the final decisions about accepting or rejecting the papers submitted for publication in the journal and shares the responsibility for the high scientific quality of the journal with the editor-in-chief.

Thus, the *Akusticheskii Zhurnal* (*Acoustical Physics*) is classed in the academic science system, and the main subject of the journal, with physical sciences. At the same time, it is clear that acoustics is interdisciplinary in character and strongly tends to applications. The letter from the editorial board to the readers that was published on the cover of the very first issue of the *Akusticheskii Zhurnal* in 1955, says "In view of the rapid development of scientific and engineering acoustics in the USSR and the need for the open publication

of the results of research and engineering works in acoustics, the Presidium of the Academy of Sciences of the USSR has acknowledged the necessity to begin in 1955 the publication of the *Akusticheskii Zhurnal* of the Academy of Sciences of the USSR." During all these years, the *Akusticheskii Zhurnal* (*Soviet Physics–Acoustics* and later *Acoustical Physics*) published both basic and applied papers. The publications that appeared in the journal in these 50 years (without taking into account the last two issues of 2004) include 80 reviews, 4802 papers, 1900 short communications, 143 letters to the editor, and 482 papers related to the chronicle and information.

The distinctive feature of the last few years is the noticeable increase in the number of papers submitted by scientists from other countries or by international teams of authors. Many publications are supported by Russian and international foundations. On the other hand, because of growing international cooperation, a considerable part of the scientific results obtained by Russian scientists is only published in foreign journals. The process of filling the "portfolio" is nonuniform and depends on many factors: the rigidity of the formulations issued by the Higher Certifying Commission and by the financial foundations about the necessity of publishing the results in Russian journals, the dynamics of the development of one or another field of acoustics in Russia and abroad, the regular availability of foreign scientific journals to Russian authors, and so on. The fluctuations in the number and scientific quality of the submitted papers require a certain flexibility in making decisions in order to provide for an acceptable quality and period of publications.

Unfortunately, the decrease in the number of subscriptions to printed periodic scientific editions is a world wide tendency. To retain the journals, one has to increase the price, which reduces the subscriptions still more. At the same time, this process is accompanied by the development of the electronic form of scientific publications. This opens up a number of fundamentally new possibilities; for example, a reader can make up an electronic journal himself by choosing articles of interest from many sources of information. For this purpose, the role of distributors should presumably be played by large organizations or publishing houses (e.g., the AIP), which own the rights for publishing the English versions of papers from a great number of journals.

Among acousticians, the conventional form of a journal is still considered to be preferable. Therefore, it is necessary to retain our printed edition for as long as possible. This requires the combined efforts of all interested individuals, including authors, reviewers, and members of the editorial board. With the understanding that we are part of the international community of acousticians, we can increase the number and improve the quality of papers submitted to the *Akusticheskii Zhurnal* (*Acoustical Physics*). Together, we can overcome such usual drawbacks as the narrowness of the scope, the small number of references to the papers from the *Akusticheskii Zhurnal* (*Acoustical Physics*) itself and from the international journals in acoustics (especially to papers that appeared within the last one or two years), the predominance of theoretical publications, and the escape of priority results to foreign journals. Both Russian and foreign specialists are interested in the existence of a strong *Akusticheskii Zhurnal* (*Acoustical Physics*). The journal can effectively progress as long as there exists the advanced Russian science that feeds it. In its turn, the journal provides for the exchange of information, without which progress in science can hardly be imagined. To contribute to these processes as much as we can is our common aim and our corporate duty.

Translated by E. Golyamina

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 105–109. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 128–132.

Detection of Moving Objects and Flows in Liquids by Ultrasonic Phase Conjugation

Yu. V. Pyl'nov[1, 2], P. Pernod[2], and V. L. Preobrazhenskii[2, 3]

[1]*Moscow State Institute of Radio Engineering, Electronics, and Automation, pr. Vernadskogo 78, Moscow, 117454 Russia*

e-mail: pylnov@yandex.ru

[2]*Institut d'Electronique, de Micro-electronique et de Nanotechnologie (IEMN-DOAE-UMR CNRS 8520), Ecole Centrale de Lille, Villeneuve d'Ascq Cedex, 59651 France*

e-mail: philippe.pernod@ec-lille.fr

[3]*Wave Research Center, Prokhorov General Physics Institute, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 119991 Russia*

e-mail: preobr@newmail.ru

Received June 23, 2004

Abstract—The possibility for the application of the method of parametric phase conjugation of ultrasonic waves in measuring the velocity of moving objects and flows is investigated. Results of experimental measurements of the Doppler frequency shift are presented for a low-frequency wave (1 MHz) generated by phase-conjugate waves (10 MHz and 11 MHz) propagating in opposite directions in the presence of a moving scatterer. The super high sensitivity of the phase of the low-frequency wave to variations in the spatial position of the scatterer is used to measure the velocity of the object. The presence of flows in the region of propagation of phase-conjugate waves returned leads to an uncompensated Doppler shift of the phase of the phase-conjugate wave at the primary radiation source. The implementation of this feature of ultrasonic phase conjugation for the detection and measurement of the flow velocities in a liquid is demonstrated experimentally. © *2005 Pleiades Publishing, Inc.*

In the last decade, ultrasonic phase conjugation has evolved into a promising and rapidly progressing field of basic and applied acoustics [1–4]. It is necessary to note that many key stages of its development were reflected in *Acoustical Physics*, from the first papers devoted to the successful implementation of the parametric phase conjugation with giant supercritical amplification [5, 6] and studies of the nonlinear processes of formation and propagation of phase-conjugate ultrasonic waves [7, 8] to the latest results in nonlinear phase-conjugation imaging [9].

In this paper, we consider some principles of using the phase-conjugation phenomenon in ultrasonic velocimetry. We present the results of studying the Doppler frequency shift of sound generated due to the combination scattering of phase-conjugate ultrasonic beams. Earlier, in [10, 11], the generation of low-frequency sound by opposite phase-conjugate waves scattered by a stationary object in a liquid was studied. A super high sensitivity of the phase of the low-frequency wave to changes in the spatial position of the scatterer was revealed. In our paper, we demonstrate the possibility of using the superhigh phase sensitivity of low-frequency sound for the Doppler detection and measurement of the velocity of moving objects. In the first section of the paper, we discuss the formation of the phase of the low-frequency sound emitted from the region of interaction of phase-conjugate beams. The second section presents the results of experimental measurements of the Doppler frequency shift for the low-frequency wave generated in the presence of a moving scatterer.

A fundamental property of phase-conjugate waves is their ability to reconstruct the phase of the primary wave at its source. The phase shifts arising in the course of the propagation of a direct wave in both homogeneous and inhomogeneous refractive media are compensated for in the course of the backward propagation of the conjugate wave. In the case of the parametric phase conjugation, the phase of the phase-conjugate wave at the source differs from the phase of the primary wave in a fixed value. The phase reconstruction is directly connected with the invariance of an acoustic field with respect to time reversal. In the presence of flows in the propagation medium, this invariance is violated. Earlier, in [2], accumulated distortions were detected in the fronts of multiply phase-conjugated waves propagating in a liquid medium with vortices. As is demonstrated in the last section of this paper, the presence of flows in the propagation region leads to an uncompensated Doppler shift in the phase of a phase-conjugate wave at the source, which may be used for measuring the flow velocity with the help of phase conjugation.

1063-7710/05/5101-0105$26.00 © 2005 Pleiades Publishing, Inc.

1. LOW-FREQUENCY ULTRASONIC GENERATION DUE TO THE INTERACTION OF PHASE-CONJUGATE WAVES IN A LIQUID

Ultrasonic generation at the difference frequency $\Omega_- = \omega_1 - \omega_2$ in the case of the combination scattering of wave beams with frequencies ω_1 and ω_2 in a liquid is described within the second order of perturbation theory by a well-known integral relation for the varying density of a medium (see [13]):

$$\rho^{(2)}(r_0, \theta, t) = \int_0^d \int_0^{2\pi} \int_0^{\infty} \frac{Q(r, z, t - R/c_0)}{R} r\,dr\,d\varphi\,dz, \quad (1)$$

where Q is the density of nonlinear sources of radiation, c_0 is the sound velocity, (r_0, θ, α) are the coordinates of the observation point, and (r, z, φ) are the coordinates of the region of nonlinear interaction. The distance R between the observation point and the interaction region in the parabolic approximation is equal to

$$R = r_0 + \frac{r^2}{2r_0} - r\sin\theta\cos(\alpha - \varphi) - z\cos\theta.$$

The specific features of the formation of the phase of low-frequency waves become clear from the model of interaction of focused high-frequency phase-conjugate beams in the region bounded by the source and its focal plane ($0 < z < d$). In the Gaussian approximation, the nonequilibrium densities in interacting ultrasonic waves are described by the expression [10]

$$\rho_{1,2}(r, z, t) = \frac{\rho_{1,2}(t, z)}{f_{1,2}(z)} \exp\left\{-\left[\frac{2}{a^2} - j\frac{q_{1,2}}{2d}\right]\frac{r^2}{f_{1,2}(z)}\right\} \times \exp[j(\omega_{1,2}t - q_{1,2}z)],$$

where $q_{1,2}$ are the wave vectors, a is the radius of the transducer aperture, and $f_{1,2}(z) = 1 - (z/d) - j(2z/q_{1,2}a^2)$. The function of nonlinear sources $Q(r, z, t)$ is equal to

$$Q = -\frac{1}{4\pi\rho_0}[(\gamma - 1) \pm 2](k_1 \mp k_2)^2 \times \rho_1\rho_2^* \exp[i\Omega_- t - i(k_1 \mp k_2)z] + \text{c.c.}$$

Here, $k_{1,2} = /q_{1,2}/$, γ is the nonlinear parameter and ρ_0 is the equilibrium density of the liquid. The upper and lower indices refer to the propagation of interacting waves in one direction and in opposite directions, respectively; $\rho_1 = \rho_1(t - z/c)$. For the case of propagation in one direction, $\rho_2 = \rho_2(t - z/c)$; for the case of opposite interacting waves, $\rho_2 = \rho_2[t + (z - d)/c]$. In the latter case, the interaction is of a nonresonance character.

As the result of integration of Eq. (1), for the nonlinear interaction of opposite waves we obtain

$$\rho^{(2)}(r_0, \theta, t) = \frac{(\gamma - 3)(k_d a)^2 \exp[i(\Omega_- t - k_d r_0)]}{16\rho_0^2 r_0} \times \frac{\exp[i(k_s - k_d\cos\theta)d]}{(k_s - k_d\cos\theta)} \times \rho_1\left(t - \frac{\rho_0}{c_0}\right)\left\{g(t) \otimes \rho_2\left(t - \frac{r_0 - d}{c_0}\right)\right\}, \quad (2)$$

where $k_d = k_2 - k_1$ and $k_s = k_2 + k_1$. The function $g(\xi)$ takes into account the degrees of focusing of ultrasonic beams and is equal to

$$g(\xi) = \frac{\exp(j\xi)}{1 + j\left(\frac{a}{2d}\right)^2 \xi}.$$

The function $g(x)$ under typical experimental conditions ($(a/2d)^2 \sim 0.01$) behaves as the Dirac δ function and almost does not distort the envelope profile of $\rho_2(t)$. As follows from Eq. (2), the low-frequency wave arising due to the nonlinear interaction of opposite phase-conjugate acoustic waves has a high phase sensitivity to the object position. At the observation angle $\theta = 0$, the phase shift is equal to $(k_s - k_d)d = 2k_2 d$, which noticeably exceeds the intrinsic phase shift of the wave of the difference frequency ($k_d d$).

2. SETUP OF THE EXPERIMENT FOR DETECTING THE DOPPLER SHIFT IN THE LOW-FREQUENCY WAVE

To radiate the initial ultrasonic wave at the frequency of 10 MHz, we used a focusing transducer with a focal distance of 30 mm. A test object in the form of a metal ball with a diameter of 300 μm was positioned in the transducer's focal region. The object could move in two directions with a step of 20 μm and could be positioned with an accuracy of 1 μm by a scanning system. After the scattering by the object, the signal arrived at the phase-conjugating device, which was a magnetoceramic cylinder with a diameter of 30 mm. The active element was pumped by an electromagnetic field at a frequency of 20 MHz. The phase-conjugate ultrasonic wave with a frequency of 10 MHz propagated back to the object, where it interacted with ultrasonic pulses emitted by the transducer at a frequency of 11 MHz. The signal of the difference frequency was detected by a transducer with a central frequency of 1 MHz. Two interaction modes were studied: the case of opposite and parallel propagation of ultrasonic beams.

In the latter case, the ultrasonic beam of 11 MHz was formed at the instant of reflection of the 10-MHz

phase-conjugate wave from the surface of the focusing transducer. Thus, two ultrasonic waves with identical wave fronts simultaneously propagated toward the object. These waves produced the difference-frequency wave due to nonlinear interaction in the focal region of the transducer.

In the case of the interaction of opposite waves, a signal at the frequency of 11 MHz was emitted with the delay necessary to provide a spatial overlapping of the pulses at the instant of reflection of the phase-conjugate wave from the object.

In the case of opposite-wave interaction, the receiver of the signal at the frequency of 1 MHz was positioned opposite the focused transducer, and the maximum of the signal at the difference frequency was detected near the receiver–transducer axis. In this mode, the super-high phase sensitivity of the difference-frequency signal to the object position was observed: while the signal frequency was determined by the wave vector $k_d = k_2 - k_1$, the signal phase corresponded to the wave vector of the nonlinear source $k_s = k_2 + k_1$.

As follows from Eq. (2), in the case of the axial position of the low-frequency receiver ($\theta = 0$), the phase shift of the ultrasonic wave is equal to

$$\Delta\varphi = (k_s - k_d)d - k_d(z - d) = k_s d - k_d z.$$

If the scattering object moves with a velocity v towards the receiver or away from it, the phase of the low-frequency signal changes according to the law $\Delta\varphi(t) = k_s v t - k_d z$, which leads to a shift of the low frequency: $\Omega_s = \Omega \pm k_s v = \Omega\left(1 \pm \frac{k_s}{k_d}\frac{v}{c}\right)$. Under the conditions of our experiment, the ratio k_S/k_d is equal to 21. As a result, the shift under study noticeably exceeds the frequency shift in common Doppler velocity meters, where it is determined by the common expression $\Omega_s = \Omega\left(1 \pm \frac{v}{c}\right)$.

The measurements of the Doppler shift were conducted for relatively low (up to 0.5 mm/s) and high (over 10 cm/s) velocities of the object motion. The repetition rate of ultrasonic pulses with a duration of 20 μs was 20 Hz. In the case of low velocities, the change in the phase of the low-frequency ultrasonic signals was detected from one pulse to another, and the Doppler frequency shift was determined by differentiating the phase with respect to time. This technique provided an opportunity to measure very low velocities of the object (from 10 μm/s). However, it could not be used to detect high velocities because of the low repetition rate of probing ultrasonic pulses. In the case of higher velocities of the object, the change in the phase within each pulse was detected. The experimental results are represented in Fig. 2. These results show that, in addition to other advantages of implementation of the phase-conjugation methods, the use of low-frequency radiation in the case of the nonlinear interaction of phase-conjugate acoustic waves provides an opportunity to considerably improve the sensitivity of Doppler velocity meters.

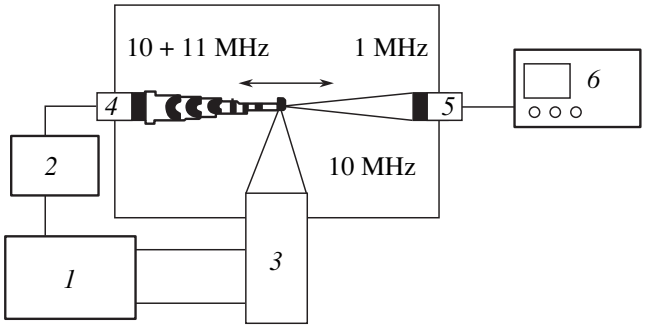


Fig. 1. Experimental setup: (*1*) pump generator, (*2*) generator of ultrasonic pulses, (*3*) phase-conjugating amplifier, (*4*) transducer with a resonance frequency of 10 MHz, (*5*) transducer with a resonance frequency of 1 MHz, and (*6*) oscilloscope.

3. THE USE OF PHASE CONJUGATION FOR DETECTING THE FLOW VELOCITY

As we have already noted above, the phenomenon of phase conjugation of acoustic waves can be also used for the detection of flow velocities in the medium of ultrasonic wave propagation. If a direct wave being transmitted through a flow moving with a velocity **v** acquires a phase shift equal to $\Delta\varphi_i = kd/(1 - v/c)$, the phase-conjugate wave has the phase shift $\Delta\varphi_c = kd/(1 + v/c)$. As a result, the phase of the received signal is equal to

$$\Delta\varphi = \Delta\varphi_i - \Delta\varphi_c = 2kd\frac{v}{c}/\left(1 - \frac{v^2}{c^2}\right).$$

Thus, the total phase shift of the received signal is proportional to the flow velocity and can be used to determine its value and direction.

In the experiment (Fig. 3), a probing ultrasonic wave was transmitted through a tube with a diameter of 5 mm. Through the tube, water was flowing with a velocity of about 100 cm/s. The tube was positioned at an angle $\alpha = 20°$ to the axis of propagation of the ultrasonic beam. The phase profile was detected in the course of the tube motion along the vertical axis. In this case, the phase of the received signal depends on the velocity of motion of the medium according to the relation

$$\Delta\varphi = 2k(r^2 - x^2)^{1/2}\frac{v}{c}\cot\alpha/\left(1 - \frac{v^2}{c^2}\cos^2\alpha\right), \qquad (3)$$

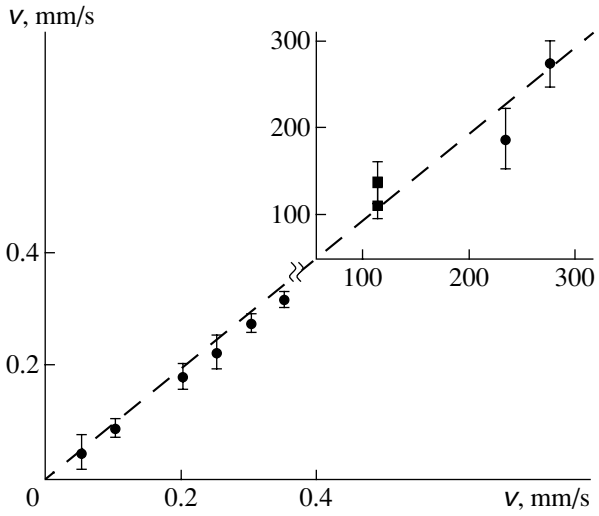


Fig. 2. Comparison between the preset velocity of the object motion (horizontal axis) and the velocity measured by the Doppler frequency shift of the low-frequency signal (vertical axis).

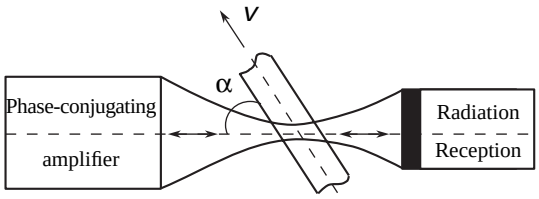


Fig. 3. Measurement of the velocity of a liquid flow by the Doppler phase shift of the phase-conjugate acoustic wave.

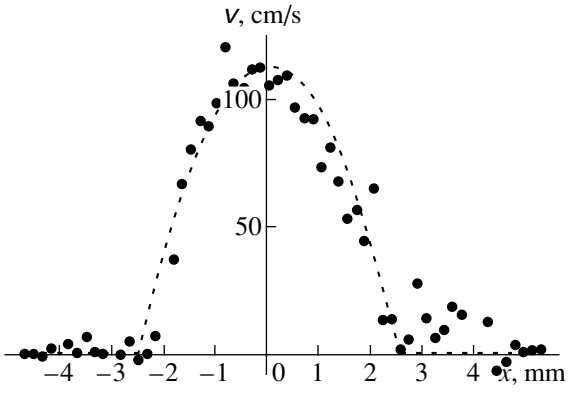


Fig. 4. Velocity distribution of the liquid flow over the cross section of a tube with a diameter of 5 mm.

where r is the tube radius and x is the distance from the tube center along the vertical axis.

Figure 4 presents the distribution of the velocity of water over the tube cross section, which was obtained by recalculation according to Eq. (3). Alternative measurements of the velocity of the water flow in the tube according to the "flow rate" demonstrated a good agreement with the aforementioned experimental data.

The shape of the velocity distribution over the tube cross section and the ratio of the average and maximum velocities of the liquid motion testify that, under the experimental conditions, the motion of the liquid flow in the tube was close to laminar.

The results presented above demonstrate that the phenomenon of phase conjugation can be used to determine the velocity of moving objects and flows in the medium of propagation of acoustic waves. In the case of measuring the velocities of moving objects, the advantages manifest themselves if one employs the low-frequency waves resulting from the nonlinear interaction of phase-conjugate waves with close frequencies that are scattered by the object. When the interaction of such waves propagating in opposite directions takes place, the wave at the difference frequency has a Doppler shift several times greater than that in the case of ordinary Doppler measurements. In the case of detecting the flows in a medium, the use of phase conjugation allows one to compensate for the phase shifts and aberrations connected with the wave propagation through the stationary part of the medium while retaining the information on the flow velocity in the phase of the phase-conjugate wave, which considerably simplifies the measurement of this velocity.

ACKNOWLEDGMENTS

This work was carried out under the program of the Russian Academy of Sciences "Coherent Acoustic Fields and Signals" and was supported by the Ministry of Education of Russian Federation and the Russian Academy of Sciences (project nos. RP-2367-MO-02 U.S. (CRDF) and VZ-010-0 (BRHE)); projects PICS (CNRS–RAS 1573), ECO–NET (08154), and Interreg 3a(56). We are grateful to Academician F.V. Bunkin for the interest taken in our work and for useful discussions.

REFERENCES

1. A. P. Brysev, L. M. Krutyanskiĭ, and V. L. Preobrazhenskiĭ, Usp. Fiz. Nauk **168**, 877 (1998) [Phys. Usp. **41**, 793 (1998)].
2. J.-L. Thomas and M. Fink, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1122 (1996).
3. A. Derode, A. Tourin, and M. Fink, J. Appl. Phys. **85**, 6343 (1999).
4. G. Montaldo, M. Tanter, and M. Fink, in *Proceedings of World Congress on Ultrasonics* (Paris, 2003), Vol. 2, p. 867.
5. A. P. Brysev, F. V. Bunkin, D. V. Vlasov, *et al.*, Akust. Zh. **34**, 986 (1988) [Sov. Phys. Acoust. **34**, 567 (1988)].
6. A. P. Brysev, F. V. Bunkin, D. V. Vlasov, *et al.*, Akust. Zh. **34**, 1120 (1988) [Sov. Phys. Acoust. **34**, 642 (1988)].

7. A. P. Brysev, F. V. Bunkin, L. M. Krutyanskiĭ, *et al.*, Akust. Zh. **44**, 738 (1998) [Acoust. Phys. **44**, 641 (1998)].

8. V. L. Preobrazhenskiĭ, Akust. Zh. **46**, 847 (2000) [Acoust. Phys. **46**, 746 (2000)].

9. A. P. Brysev, L. M. Krutyanskiĭ, Ph. Pernod, and V. L. Preobrazhenskiĭ, Akust. Zh. **50** (6), 725 (2004) [Acoust. Phys. **50** (6), 623 (2004)].

10. Yu. V. Pylnov, V. L. Preobrazhensky, and Ph. Pernod, Acta Acust. **89** (6), 942 (2003).

11. Ph. Pernod, V. Preobrazhensky, and Yu. V. Pylnov, in *Proceedings of World Congress on Ultrasonics* (Paris, 2003), Vol. 2, p. 855.

12. P. Poux and M. Fink, in *Proceedings of 4th French Congress on Acoustics*, Ed. by G. Canevet (Teknea, Marseille, 1997), Vol. 1, p. 951.

13. O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Nauka, Moscow, 1975; Consultants Bureau, New York, 1977).

Translated by M. Lyamshev

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 11–23. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 24–36.

Radiation from Uniformly Moving Sources (Vavilov–Cherenkov Effect, Transition Radiation, and Some Other Phenomena)[1]

V. L. Ginzburg

Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 117924 Russia

Received September 9, 2004

Abstract—The radiation produced by uniformly moving sources (the Vavilov–Cherenkov effect, the transition radiation, and some other phenomena) is discussed. This area of physical research originated in the Lebedev Physical Institute of the Russian Academy of Sciences and now represents an integral part of modern physics.

1. INTRODUCTION

During my long life, I have worked in many fields of science, which is typical of theoretical physicists. For the topic of this lecture, I could choose between the theory of superconductivity, astrophysics of cosmic rays, and radiation of uniformly moving sources. I choose the latter for two reasons. The first is that I love this area of research. Of course, the word "love" is not often used in scientific literature, but this is only a tribute to conventional style. In fact, in science, as in everyday life, we all love certain things and dislike some others. I love the problems concerned with the radiation of uniformly moving sources, because my first scientific results are related to this subject and remind me of my youth. The second reason for choosing this lecture topic is that the radiation from uniformly moving sources represents a basically Russian and, in addition, academic area of research. Indeed, the brightest phenomenon in this area, namely, the Vavilov–Cherenkov (V–Ch) effect, was discovered by S.I. Vavilov and P.A. Cherenkov in 1934 [1, 2]. The effect was interpreted by I.E. Tamm and I.M. Frank in 1937 [3]. The transition radiation was first considered by I.M. Frank and myself in 1945 [4]. All the aforementioned authors worked at the Lebedev Physical Institute and all of them were members of the Academy of Sciences of the USSR. In 1958, I.E. Tamm, I.M. Frank, and P.A. Cherenkov received the Nobel Prize in Physics for the discovery and interpretation of the V–Ch effect (Vavilov died in 1951 at less than sixty years of age, and the Nobel Prize could not be given posthumously).

2. THE VAVILOV–CHERENKOV EFFECT

In the framework of its own, somewhat narrowed interpretation, the V–Ch effect is as follows: an electric charge (e.g., an electron) moving in a medium with a constant velocity **v** emits electromagnetic waves (light) with a continuous spectrum and with a specific angular distribution. Radiation at a cyclic frequency ω occurs only if the velocity of the charge **v** exceeds the phase velocity of light in the transparent medium under consideration, $v_{ph} = c/n(\omega)$:

$$v > \frac{c}{n(\omega)}, \tag{1}$$

where $n(\omega)$ is the refraction index for light at the frequency ω in the medium and c is the velocity of light in vacuum. The aforementioned specificity of the angular distribution of the radiation consists in that the wave vector of the emitted waves **k** and the velocity **v** make an angle θ_0 characterized by the formula

$$\cos\theta_0 = \frac{c}{n(\omega)v}. \tag{2}$$

Results (1) and (2) can be obtained by applying the Huygens principle: every point on the path of a charge moving uniformly along a straight line with a velocity v represents a source of a spherical wave, which is emitted at the instant of the charge passage through this point (Fig. 1). Under condition (1), the spheres have a common envelope in the form of a cone whose vertex coincides with the instantaneous position of the charge, while the angle θ_0 is determined by Eq. (2).

If the dispersion, i.e., the dependence of n on ω, is ignored, the angle θ_0 is the same for all frequencies ω, and the radiation has a sharp wave front forming a cone with an angle of $\pi - 2\theta_0$ and with the charge (source) at its vertex (see Fig. 1). This cone is similar to the Mach cone, which characterizes a shock wave accompanying supersonic motion of a source (bullet, missile, airplane, or rocket) in air or in another medium. In this case, the role of the phase velocity of light $v_{ph} = c/n$ involved in expressions (1) and (2) is played by the velocity of the shock wave or the velocity of sound u. Since the dispersion of sound, i.e., the dependence of its velocity u on

[1] This lecture was given by V.L. Ginzburg upon receiving the Lomonosov Large Gold Medal, the highest prize awarded by the Russian Academy of Sciences. The lecture was published in *Physics–Uspekhi* **39** (10), 973 (1996).

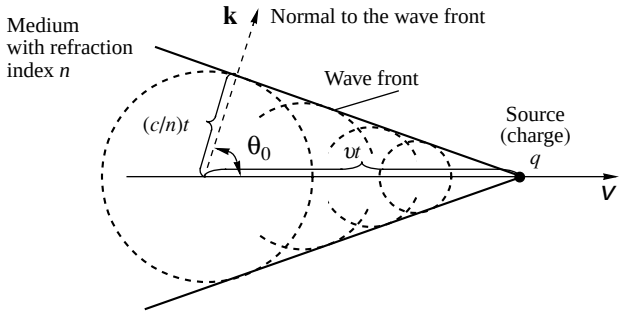


Fig. 1. Formation of the Vavilov–Cherenkov effect: $(c/n)t$ is the path length traveled by light within the time t and $\upsilon t = [c/(n\cos\theta_0)]t$ is the path length traveled by the charge (source) within the same time.

frequency, is usually very small, the hydrodynamic (acoustic) wave front at the Mach cone surface is sharp and is often observable (e.g., at the passage of a supersonic airplane).

Thus, V–Ch radiation represents an electrodynamic (optical) analog of the well-known (since the nineteenth century) acoustic phenomenon. Why, then, was V–Ch radiation not discovered and explained only as late as only about 60 years ago? Of course, this could have happened earlier, but, on the whole, the delay is not accidental. First, the observation of the V–Ch effect in a relatively pure form requires a beam of relativistic or nearly relativistic charged particles. Such beams were obtained in only the 1930s (when the first accelerators were built). Second, in electrodynamics (unlike hydrodynamics and acoustics), the motion of sources (charges) is primarily and most commonly considered in vacuum. Since the velocity of particles v is always smaller than the velocity of light $c = 3 \times 10^{10}$ cm s^{-1} (here, we do not consider the hypothetical and, most likely, nonexistent supraluminal particles, i.e., tachyons), the V–Ch effect in vacuum is impossible. Though here some reservations are appropriate (see, e.g., Sect. 9 in [5] and [6, 7]), on the whole, the existence of the former statement that "a uniformly moving charge does not emit radiation" is quite understandable.

Presumably, this dogma did not allow one to predict the V–Ch effect earlier. Actually, however, such a prediction was made by the well-known English physicist Heaviside in 1888 [8], but, at that time, even the electron had not been discovered and no fast particles moving in a dielectric could be imagined to exist in reality. Therefore, Heaviside's idea was forgotten and was not again brought to light until as late as 1974 [9, 10]. Another precursor of the Tamm and Frank theory was the calculation performed by the well-known German physicist Sommerfeld [11], but Tamm and Frank became acquainted with it only after the termination of their work [3]. In 1904, Sommerfeld considered a uniform motion of a charge in a vacuum and made the conclusion that, at a supraluminal velocity $v > c$, the charge emits radiation. However, from the special relativity theory, which appeared within a year (in 1905), it followed that the motion of a charge with a velocity higher than c is impossible, and the work by Sommerfeld was forgotten as well.[2] Within the next 30 years, neither Sommerfeld nor any other physicists tried to consider the motion of a charge in a medium instead of the vacuum.

Precisely this idea was put forward by I.E. Tamm and I.M. Frank [3]: they calculated the radiation from a charge q moving with a constant velocity $\mathbf{v}$ in a medium with a refraction index $n(\omega)$. As a result, they formally derived expression (2) and obtained the radiation intensity (power) per unit time (i.e., within a path length equal to v):

$$\frac{dW}{dt} = \frac{q^2 v}{c} \int\limits_{c/[n(\omega)v] \le 1} \left(1 - \frac{c^2}{n^2(\omega)v^2}\right)\omega d\omega. \qquad (3)$$

Here, the integration is performed over all frequencies satisfying condition (1). Tamm and Frank sent a reprint of their paper [3] to Sommerfeld and, in response, received a letter dated May 8, 1937, which arrived through Austria (fascists were already in power, and it was difficult to write a letter directly to the USSR). In this letter, Sommerfeld wrote: "I never thought that my calculations of 1903 may find any physical application. This case proves that the mathematical part of the theory experiences a change of physical concepts."[3]

The prehistory of the discovery of the V–Ch effect is described in more detail in the book written by Frank [14]. The development of the theory has been considered above. As for the experiment, the V–Ch radiation had actually been observed by Pierre and Marie Curie in bottles with radium salt solutions. Today, the blue glow of water, which mainly represents the V–Ch radiation, can be observed by excursionists when they are shown a nuclear reactor immersed in a water tank. The radiation from fluids irradiated with gamma rays was studied by the French scientist Mallet in 1926–1929. However, before Vavilov and Cherenkov, no one understood that this phenomenon was related to a new effect rather than to some kind of luminescence under the effect of gamma rays.

---

[2] Note that, from relativity theory, it follows (without considering tachyons) that the velocity c is limiting for a single charge (when $v \longrightarrow c$, the mass of the particle $m_0/\sqrt{1 - v^2/c^2}$ tends to infinity). However, a source of radiation (for example, a source consisting of many particles) may have any velocity (see [5–7]). I do not discuss this problem in this lecture, although it is of a certain interest.

[3] This letter is fully reproduced in *Recollections of I.E. Tamm* (see p. 120 in [12]). Sommerfeld also mentioned that, as a foreign member of the Academy of Sciences of the USSR, he received some of the Russian academic literature. Presumably, he was referring to the *Doklady Akademii Nauk SSSR* (*Doklady Physics*). It is a pity that today's foreign members of the Russian Academy of Sciences receive nothing [13].

Experiments suggested by Vavilov and carried out by Cherenkov began with the observation of the luminescence of uranyl salt solutions under the effect of gamma rays. The experiments used an original method of measurement that was developed by Vavilov and his coworkers on the basis of using a dark-adapted human eye [14, 15]. By accident, Cherenkov discovered that the fluid (sulfuric acid) glows even in the absence of salt dissolved in it. He decided that his thesis had failed [15]. However, Vavilov understood that the experiments revealed a glow of a nature different from luminescence. Vavilov and Cherenkov continued their measurements and finally obtained enough evidence to report on the discovery of a new phenomenon [1, 2]. Vavilov indicated [2] that the glow observed in the experiment was not caused by gamma rays but was related to the Compton electrons that appeared in the fluid as a result of their knocking-out by gamma rays. The subsequent observations by Cherenkov [16] were performed with the participation of Vavilov and Frank [14, 15]. They revealed some properties of the new radiation that allowed Tamm and Frank to determine its nature [3].

From the aforesaid, it is clear that Vavilov is the coauthor of the discovery of the V–Ch effect, and only the name "Vavilov–Cherenkov" should be used for it. I stress this point because, in our (Soviet) literature, other opinions can be found, ones which are considered wrong by all physicists knowing the actual facts (see [14, 15, 17, 18]). As for the name "Cherenkov effect," which is always used abroad and sometimes in Russia—it is a consequence of the actions of Vavilov himself: he first published only a small letter reporting on the effect [2] and then sent to the *Physical Review* a large paper describing the V–Ch effect in detail [19], but this paper was signed by Cherenkov alone.[4] I do not know why Vavilov did such a thing. Possibly, because of his noble nature, he did not want to eclipse his student. Unfortunately, Vavilov suffered from various kinds of malicious attacks as a physicist, as a person, and as a scientific administrator and President of the Academy of Sciences of the USSR. I believe that all this criticism was groundless, and this opinion has already been expressed by me in my previous publications (see pp. 391, 393 in [20] and also [21]).

The V–Ch effect has found wide application in physics (here, I do not mention its significance for the electrodynamics of continuous media and for physics as a whole). On the basis of the V–Ch effect, one can determine the velocity of a particle v by measuring the angle θ_0 (see Eq. (2)) or, from inequality (1), directly (in the absence of the effect) conclude that $v < c/n(\omega)$ (evidently, the refraction index $n(\omega)$ in a transparent medium may and should be considered to be known). In addition, since the radiation intensity is proportional to the square of the charge q of a particle (see Eq. (3)), one can easily distinguish particles with the elementary charge e (electrons, protons, etc.) from nuclei with charges Ze (where Z is the ordinal number of an element). Even for a helium nucleus ($Z = 2$), the radiation intensity is four times as high as that for hydrogen isotopes ($Z = 1$); for an iron nucleus ($Z = 26$), the intensity is 676 times higher that for protons with the same velocity. The so-called Cherenkov counters are widely used in accelerators and in high-energy physics as a whole [22, 23]. Specifically, the V–Ch effect is used in studying cosmic rays (the V–Ch radiation from a shower in the atmosphere) and in systems designed for observing high-energy neutrinos.

The theory of V–Ch radiation cannot be fully described in the framework of this lecture (for details, see [5–7, 14, 22, 24]), and I will dwell only on several problems that are subjects of my own former investigations.

In 1940, Mandel'shtam, acting as an official opponent to Cherenkov's doctoral thesis, noted that the V–Ch effect should also be observed when a charge (source) moves not in a continuous medium but in a thin empty channel running through this medium. Physically, the idea is that the V–Ch radiation is formed not only on the very path of the charge but also near the path, at a distance of about the wavelength of the emitted light $\lambda = 2\pi c/[n(\omega)\omega]$. The corresponding radiation intensity was calculated by Frank and myself [25]. Naturally, this intensity decreases with increasing radius r of the empty channel, along the axis of which the charge moves. If we have $\sqrt{1 - v^2/c^2} \sim 1$, then, for $r/\lambda \lesssim 0.01$ (in optics, this means that $r \lesssim 5 \times 10^{-7}$ cm), the radiation is practically the same as that in the absence of the channel. Qualitatively, a similar situation occurs when the channel is replaced by a gap or when the charge moves near the boundary of a medium (a dielectric). This consideration is important because, in the course of the motion of the charge in a medium, its energy loss due to V–Ch radiation is relatively small; the predominant factor is the ionization loss, which is localized in the immediate vicinity of the trajectory. Therefore, in the case of motion in a channel, in a gap, or near a medium, the ionization loss is absent though the V–Ch radiation persists. For charges, this fact is important but not critical, while for the observation of the Doppler effect in a medium, where the motion of excited atoms is involved, the whole phenomenon can be observed only with the use of channels or gaps; otherwise, the atom breaks up. However, the Doppler effect may be observed (and was observed) in the case of the motion in a rarefied medium, in particular, in plasma.

Incidentlly, the analysis of the problem of radiation in the case of motion near a medium was used by me in

[4] Note that this paper [19] was initially sent to *Nature* but was rejected. This shows that the V–Ch effect seemed to be rather nontrivial at that time.

discussing various possibilities of microwave generation [26–28].

Now, I consider the methods of calculating the intensity of V–Ch radiation. Tamm and Frank [3] obtained expression (3) by solving the electrodynamics equations in a medium and defining the radiation intensity as the Poynting vector flux through a cylindrical surface surrounding the trajectory of the charge. Another method of calculation consists in the determination (on the basis of the same equations) of the force that decelerates the moving charge: the work of this force in a transparent medium is equal to the radiation energy given by Eq. (3). Such calculations were performed, for example, by Fermi [29], and they can be found in the book by Landau and Lifshits [30] (see § 115). Finally, there is a third method for calculating the same intensity (power) given by Eq. (3). It consists of calculating the energy of the electromagnetic field generated by the charge per unit time [31].

For this purpose, it is convenient to use the so-called Hamiltonian method. For a homogeneous isotropic stationary medium, it uses a series expansion of the vector potential **A** of the field (for more details, see, e.g., [5]):

$$\begin{cases} \mathbf{A}(\mathbf{r}, t) = \sum\limits_{\lambda,\, i\,=\,1,\,2} q_{\lambda i}(t)\mathbf{A}_{\lambda i}(\mathbf{r}); \\ \mathbf{A}_{\lambda 1} = \mathbf{e}_\lambda \sqrt{8\pi}\dfrac{c}{n}\cos(\mathbf{k}_\lambda \mathbf{r}), \\ \mathbf{A}_{\lambda 2} = \mathbf{e}_\lambda \sqrt{8\pi}\dfrac{c}{n}\sin(\mathbf{k}_\lambda \mathbf{r}), \end{cases} \tag{4}$$

where $\mathbf{e}_\lambda$ is the polarization vector ($e_\lambda = 1$) and $n = \sqrt{\varepsilon}$ is the refraction index (ε is the permittivity of the medium, and the latter is assumed to be nonmagnetic for simplicity). The transverse electromagnetic field under consideration has the form

$$\mathbf{E}_{\text{tr}} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{H} = \text{rot}\mathbf{A},$$

and the energy of this field is

$$\mathcal{H}_{\text{tr}} = \int \frac{\varepsilon E_{\text{tr}}^2 + H^2}{8\pi} dV = \frac{1}{2}\sum_{\lambda,\, i\,=\,1,\,2} (p_{\lambda i}^2 + \omega_\lambda^2 q_{\lambda i}^2), \tag{5}$$

where

$$p_{\lambda i} = \frac{dq_{\lambda i}}{dt}, \quad \omega_\lambda^2 = \frac{c^2}{\varepsilon}k_\lambda^2 \equiv \frac{c^2}{n^2}k_\lambda^2. \tag{6}$$

The field equation has the form

$$\Delta \mathbf{A} - \frac{\varepsilon}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{4\pi}{c}\mathbf{j};$$

for a point charge q moving with a velocity $\mathbf{v}$, the current density is $\mathbf{j} = q\mathbf{v}\delta[\mathbf{r} - \mathbf{r}_q(t)]$, where $\mathbf{r}_q(t)$ is the radius vector of the charge and δ is a delta function. With the substitution of expansion (4), the field equation takes the form

$$\begin{cases} \dfrac{d^2 q_{\lambda 1}}{dt^2} + \omega_\lambda^2 q_{\lambda 1} = \sqrt{8\pi}\dfrac{c}{n}(\mathbf{e}_\lambda \mathbf{v})\cos(\mathbf{k}_\lambda \mathbf{r}_q), \\ \dfrac{d^2 q_{\lambda 2}}{dt^2} + \omega_\lambda^2 q_{\lambda 2} = \sqrt{8\pi}\dfrac{c}{n}(\mathbf{e}_\lambda \mathbf{v})\sin(\mathbf{k}_\lambda \mathbf{r}_q). \end{cases} \tag{7}$$

Thus, the field equations are reduced to Eqs. (7) for the "field oscillators" $q_{\lambda i}(t)$. Integrating these equations and substituting the solution into Eq. (5), we obtain the field energy as the sum of the energies of all oscillators. For a charge uniformly moving along a straight line, we have $\mathbf{r}_q(t) = \mathbf{v}t$, and Eqs. (6) can be easily integrated, because they are equations for an oscillator oscillating under a harmonic force proportional to $\cos(\mathbf{k}_\lambda \mathbf{v}t)$ or $\sin(\mathbf{k}_\lambda \mathbf{v}t)$, i.e., oscillating with a frequency

$$\omega = \mathbf{k}_\lambda \mathbf{v} = k_\lambda v\cos\theta = \frac{\omega_\lambda n v}{c}\cos\theta. \tag{8}$$

At $\omega = \omega_\lambda$, a resonance takes place and the amplitudes $q_{\lambda i}$ increase with time; i.e., radiation is observed. Evidently, in a vacuum where $n = 1$, the frequency ω is always smaller than ω_λ (provided that $v < c$). This means that a charge moving uniformly in a vacuum does not emit any radiation. In a medium, however, the resonance (and, hence, the radiation) is possible. According to Eq. (8), the resonance condition is $(nv/c)\cos\theta = 1$, i.e., the V–Ch radiation condition (see Eq. (2)). The substitution of the solution for $q_{\lambda i}(t)$ into Eq. (5) yields the expression $\mathcal{H}_{\text{tr}} = (dW/dt)t$, where dW/dt is determined by Eq. (3).

Thus, in the case under consideration, calculation by the Hamiltonian method is simple and self-evident. The character of this lecture allows me to note that precisely this simplicity, which I discovered by accident, encouraged me to switch to theoretical physics (I graduated from the Moscow State University in 1938 as an experimentalist in optics and believed that I should not work in theoretical physics because of the lack of mathematical talent). Of course, I considered the Hamiltonian method not with the aim of making the above comment. It is important that, unlike the two other methods mentioned above, the Hamiltonian method for calculating the radiation energy allows for an almost trivial generalization to the case of an anisotropic medium, i.e., to noncubic crystals or to a plasma in a magnetic field. In an anisotropic medium, the field should be decomposed into normal waves, which may propagate in the corresponding medium (in an isotropic medium, as in vacuum, degeneration takes place and the normal waves are reduced to waves $\mathbf{A}_{\lambda i}$ given by Eqs. (4)). Thus, one can easily consider the V–Ch effect in an anisotropic medium, the simplest example of which is a uniaxial crystal [32]. In this case, the V–Ch radiation forms two cones, which, in the general case, are not circular and have different polarizations (directions of the electric

field in the waves). Experimental studies of the V–Ch effect in crystals were performed, in particular, by Zrelov [22].

Numerous publications in addition to those cited are concerned with different aspects of the theory of V–Ch radiation. They include the generalization to magnetic media, a detailed analysis of radiation in crystals, the role of boundaries, etc. (see [5–7, 14, 22, 33, 34] and the literature cited there). I specially note the study of the role of absorption [29, 35] and the consideration of V–Ch radiation not from charges but from various dipoles and multipoles (see [5–7, 14] where references to original publications are given). The problem of the V–Ch radiation from multipoles has not been completely investigated [5–7]. Presumably, this can be explained by the fact that the known particles have rather small magnetic moments (not to mention other multipoles), and the radiation associated with these moments is also very weak and of no practical interest. As for the radiation from a magnetic charge (a monopole), it should be considerable, but such monopoles have never been observed; possibly, they do not exist in nature.

For reasons of space, it is impossible to dwell here on the problems listed above and on all the experiments using V–Ch radiation (see [22, 23]). However, the quantum interpretation of the V–Ch effect seems to be worth discussing.

3. THE QUANTUM THEORY OF THE VAVILOV–CHERENKOV EFFECT

The classical theory of the V–Ch effect, which was discussed above, is sufficiently accurate in the optical part of the spectrum. Nevertheless, proceeding from methodical considerations, it is expedient to consider more closely the quantum theory of the effect [36] (see also [5–7], 14]).

How should the absence of radiation from a charge (or some other source possessing no eigenfrequency) uniformly moving in vacuum be explained in quantum terms? For this purpose, it is sufficient to use the energy and momentum conservation laws in application to the emission of a photon by a particle:

$$\begin{cases} E_0 = E_1 + \hbar\omega, \quad E_{0,1} = \sqrt{m^2c^4 + c^2p_{0,1}^2}, \\ \mathbf{p}_0 = \mathbf{p}_1 + \hbar\mathbf{k}, \quad k = \dfrac{\omega}{c}, \quad \mathbf{p}_{0,1} = \dfrac{m\mathbf{v}_{0,1}}{\sqrt{1 - v_{0,1}^2/c^2}}, \end{cases} \tag{9}$$

where $E_{0,1}$ and $\mathbf{p}_{0,1}$ are the energy and momentum of a charge with a mass at rest m before the emission (subscript 0) and after the emission (subscript 1) of a photon with an energy $\hbar\omega$ and momentum $\hbar\mathbf{k} = (\hbar\omega/c)(\mathbf{k}/k)$. One can easily verify that Eqs. (9) have no solution for $v < c$ (with $\omega > 0$); i.e., the photon emission is impossible (see Eq. (11) with $n = 1$).

To consider the problem of radiation from a source in a medium, it is necessary to know only: what the energy and momentum of the radiation are in this case, because the energy of a particle $E = \sqrt{m^2c^4 + c^2p^2}$ does not change in the medium. This question is not that simple (see Sect. 13 in [5]), but it can be solved rather simply (and correctly) at the intuitive level. Indeed, in a stationary and time-invariant medium, its presence does not affect the frequency ω, and the wavelength is $\lambda = \lambda_0/n$, where $\lambda_0 = 2\pi c/\omega$ is the wavelength in vacuum. The wave number is $k = 2\pi/\lambda = \hbar\omega n/c$. Taking this into account, instead of Eqs. (9), we write

$$\begin{cases} E_0 = E_1 + \hbar\omega, \quad E_{0,1} = \sqrt{m^2c^4 + c^2p_{0,1}^2}, \\ \mathbf{p}_0 = \mathbf{p}_1 + \hbar\mathbf{k}, \\ k = \dfrac{\hbar\omega n(\omega)}{c}, \quad \mathbf{p}_{0,1} = \dfrac{m\mathbf{v}_{0,1}}{\sqrt{1 - v_{0,1}^2/c^2}}. \end{cases} \tag{10}$$

Solving these equations with respect to ω and θ_0, where θ_0 is the angle between $\mathbf{v}_0$ and $\mathbf{k}$, we obtain

$$\cos\theta_0 = \frac{c}{n(\omega)v_0}\left[1 + \frac{\hbar\omega(n^2 - 1)}{2mc^2}\sqrt{1 - \frac{v_0^2}{c^2}}\right], \tag{11}$$

$$\hbar\omega = \frac{2(mc/n)(v_0\cos\theta_0 - c/n)}{(1 - 1/n^2)\sqrt{1 - v_0^2/c^2}}. \tag{12}$$

Under the condition

$$\frac{\hbar\omega}{mc^2} \ll 1 \tag{13}$$

(or under a more accurate inequality evident from Eq. (11)), expression (11) is transformed into classical expression (2). This result should be expected, because condition (13) evidently corresponds to the classical limit (it is always valid when $\hbar \longrightarrow 0$). In the classical limit, the recoil (the change in the momentum $\mathbf{p}_0$ of the particle) due to the emission of a photon in the medium with a momentum $\hbar\mathbf{k}$ is ignored. As was mentioned above, from Eq. (12) it follows that emission ($\omega > 0$) is possible only when $v_0 > c/n$ (because $\cos\theta_0 \le 1$). In the classical limit, when the result (see Eq. (2)) does not depend on $\hbar$, the quantum calculation is only of methodical significance; it may be convenient but is not obligatory. This corresponds to the real situation, and the conservation laws can be formulated in the classical region as well; it is only necessary to take into account the relation between the emitted electromagnetic energy $\mathcal{H}_{tr}$ and the momentum of the radiation. The corresponding simple calculations are given in [5–7].

Evidently, the radiation intensity can also be calculated in quantum terms by generalizing Eq. (3) [36].

In the optical region, which is the only one dealt with in the applications of the V–Ch effect, even for electrons the ratio $\hbar\omega/mc^2 \sim 10^{-5}$ is small; i.e., the quantum corrections are insignificant. In light of this, Landau, when he became aware of my work [36] published in 1940, noted that it is of no interest (see p. 380 in [20]). From the aforesaid, it is clear that this conclusion was justified, and this was characteristic of Landau: his critical comments were usually correct. However, a different approach or method of obtaining a known result proves to be useful in applications to other problems. Such an example concerning different ways of calculating the V–Ch radiation power given by Eq. (3) was mentioned above. It was found that a similar situation arises in the case of applying the conservation laws to the analysis of the radiation in a medium. Precisely the application of conservation laws proves to be fruitful in studying the Doppler effect in a medium.

4. THE DOPPLER EFFECT IN A MEDIUM

The sources discussed above (specifically, charges) possess no eigenfrequency. Another important case is a source without a charge or any time-invariable multipole moment but with an eigenfrequency ω_0. A classical example is an oscillator, and a quantum example is an atom, which, at a certain transition, radiates a frequency ω_{00} (this frequency refers to the frame of reference in which the source is at rest).

If such a source moves with a constant velocity **v** (in the laboratory frame of reference) in a vacuum, the frequency of the waves produced by it is estimated in the laboratory frame of reference by the formula

$$\omega(\theta) = \frac{\omega_{00}\sqrt{1 - v^2/c^2}}{1 - (v/c)\cos\theta} = \frac{\omega_0}{1 - (v/c)\cos\theta}, \qquad (14)$$

where θ is the angle between the wave vector **k** (the observation direction) and the velocity **v**; the frequency ω_0 in Eq. (14) represents the oscillation frequency in the laboratory frame of reference. The change in the frequency of waves produced by a moving source is called the Doppler effect. This effect also occurs in acoustics, as well as for waves of any nature.

Now, let us consider a transparent medium (with a refraction index $n(\omega)$), which is at rest in the same laboratory frame of reference, and an oscillator or an atom (molecule) moving in it. The fact that a source moving through a continuous medium may fail is insignificant, because a channel or a gap in the medium can be used (see above).

In the presence of the medium, Eq. (14) should be replaced by the expression [37, 14]

$$\omega(\theta) = \frac{\omega_{00}\sqrt{1 - v^2/c^2}}{|1 - (v/c)n(\omega)\cos\theta|} = \frac{\omega_0}{|1 - (v/c)n(\omega)\cos\theta|}. \qquad (15)$$

This formula can be obtained using the following general rule: by replacing the velocity of light in vacuum by the phase velocity in the medium $c/n(\omega)$ (in the radicand of $\sqrt{1 - v^2/c^2}$, the quantity c should not be replaced by c/n, because this root is associated with the time dilation for the moving source and is unrelated to the radiation process). Expression (15) can also be derived formally by solving the field equations for a moving source. In this expression, the appearance of the absolute value is nontrivial (its necessity is evident from the requirement that the frequency should be positive). If the motion is subluminal (i.e., $v < c/n$) or if it is supraluminal but occurs outside the cone defined by Eq. (2), i.e., under the condition that

$$\frac{v}{c}n(\omega)\cos\theta < 1, \qquad (16)$$

the conventional normal Doppler effect takes place. Note that, in this case, the so-called complex Doppler effect, which is caused by the dispersion, i.e., by the dependence of n on ω, is also possible [37, 14].

If the motion is supraluminal (condition (1) is satisfied), then, in the angular region determined by the formula

$$\frac{v}{c}n(\omega)\cos\theta > 1, \qquad (17)$$

Eq. (15) without the absolute value sign leads to negative frequency values. The radiation in region (17), i.e., inside the cone defined by Eq. (2) (often called the Cherenkov cone (Fig. 2)), is called the anomalous Doppler effect. With allowance for the dispersion, the whole picture proves to be complicated (every frequency has its own cone and, if the dependence of n on ω is nonmonotone, several cones). Here, we limit our consideration to the case without dispersion: $n(\omega) = n = \text{const}$. Then, according to Eq. (15), on the Cherenkov cone surface where $(v/c)n\cos\theta = (v/c)n\cos\theta_0 = 1$, the frequency is $\omega \longrightarrow \infty$, and this occurs on both sides of the cone surface (at $\theta \longrightarrow \theta_0$). From Eq. (15), no other inferences can be drawn, and the difference between the normal and anomalous Doppler effect does not seem to be significant.

However, it was found that the quantum approach (or, more precisely, the use of the energy and momentum conservation laws) reveals a very important feature of the anomalous Doppler effect [38, 5–7, 14]. Let us assume that the source is a "system" (atom) with two levels: the lower level 0 and the upper level 1 (Fig. 3).

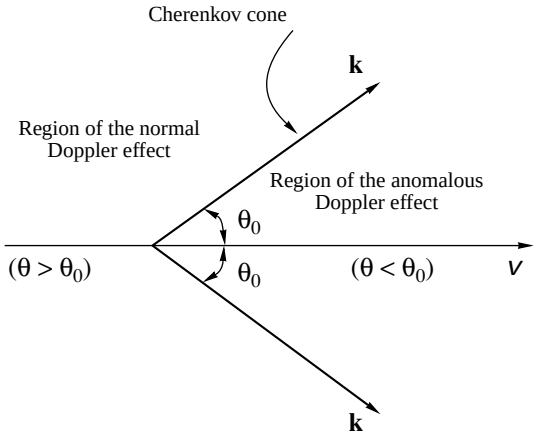


Fig. 2. Regions of the normal ($\theta > \theta_0$) and anomalous ($\theta < \theta_0$) Doppler effects.

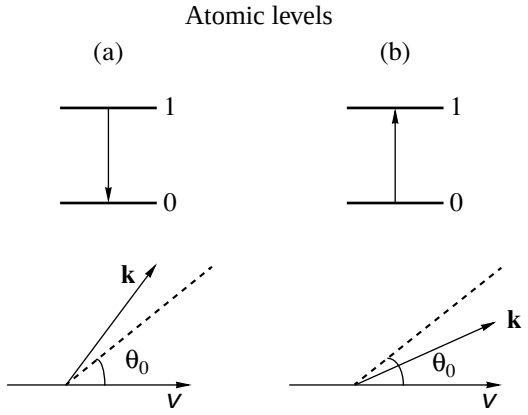


Fig. 3. Transitions between atomic levels 0 and 1 in the cases of the (a) normal and (b) anomalous Doppler effects.

Then, using conservation laws of type (10), one only has to modify the expression for the energy of the source by taking into account the presence of the internal degrees of freedom (levels). This energy has the form

$$E_{0,1} = \sqrt{(m + m_{0,1})^2 c^4 + c^2 p_{0,1}^2}, \qquad (18)$$

where $(m + m_0)c^2 = mc^2 + W_0$ is the total energy of the system (atom) in the lower state 0 and $(m + m_1)c^2 = mc^2 + W_1$ is the corresponding energy in the upper state 1. Since $W_1 > W_0$, in the case of the transition $1 \longrightarrow 0$, the atom at rest emits radiation with the frequency $\omega_{00} = (W_1 - W_0)/\hbar$.

Using the conservation laws in classical limit (13), we arrive at Eq. (14) and, in the case of an exact calculation [38], to a somewhat more complex expression containing terms on the order of $\hbar\omega/mc^2$. However, not the quantum corrections but the following unexpected situation proves to be significant. Tracing the signs (this is simple algebra), one can easily see that, in the region of the normal Doppler effect, the atom, as in vacuum, performs a transition from the upper level 1 to the lower level 0 (the direction of the transition is determined from the requirement that the energy of the emitted photon, $\hbar\omega$, be positive, i.e., from the condition $\omega > 0$). By contrast, in the region of the anomalous Doppler effect, the emission of a photon is accompanied by the excitation of the atom: it performs the transition from level 0 to level 1 (Fig. 3). In this case, the energy is taken from the kinetic energy of the translational motion.

Thus, in the case of supraluminal motion ($v > c/n$), which is the only case allowing the anomalous Doppler effect, the initially nonexcited atom (in the lower state 0) becomes excited (changes to level 1), simultaneously emitting a photon within the Cherenkov cone. The excited atom emits radiation (upon the transition $1 \longrightarrow 0$) outside the Cherenkov cone, i.e., at the angles $\theta > \theta_0$. As a result, in the course of the supraluminal motion, the atom is continuously excited and emits radiation. For the classical oscillator model, this means that the oscillator is excited all the time. The anomalous Doppler effect is quite important for plasma physics. On the whole, in plasma, the V–Ch effect and the notions and analogies related to it play an important role. This fact was stressed by Tamm in his Nobel lecture [17]. There, he also put forward the assumption that the acoustic analog of the anomalous Doppler effect known in optics plays an important part in the analysis of vibrations that accompany the supersonic motion of an airplane (the so-called flutter).

I think that, without the quantum consideration, it would be difficult to reveal the aforementioned feature of the anomalous Doppler effect [38] (more precisely, as follows from the above, not the quantum approach but the application of the conservation laws is important). The testing of the result and further progress is made possible by means of a classical or quantum calculation of the radiation response for a source moving in a medium. Specifically, for an oscillator moving in a medium, one can determine the effect of the radiation force on the oscillations of the oscillator (see [39], [5], ch. 7). The wave radiation in the region outside the Cherenkov cone (i.e., in the case of the normal Doppler effect) was found to slow down the oscillations. By contrast, the radiation directed inside the Cherenkov cone, i.e., corresponding to the anomalous Doppler effect, enhances the oscillations of the oscillator, i.e., excites it. This result agrees well with the above quantum consideration.

Note that a number of publications developing the cited work [39] and also some other publications in this area of research belong to B.E. Nemtsov [40], the well-known governor of the Nizhni Novgorod region and a former talented theoretical physicist.

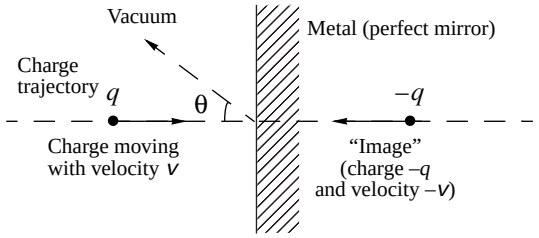


Fig. 4. Transition radiation from a charge q crossing the vacuum–metal boundary.

The above consideration also clarifies the mechanism of excitation of uniformly accelerated "detectors" [41, 7]. As is known, the latter problem has been much discussed in the literature (for references, see [41]) in connection with studies of the radiation from black holes and uniformly accelerated systems (acceleration radiation).

5. TRANSITION RADIATION AT A BOUNDARY BETWEEN TWO MEDIA

In the course of uniform rectilinear motion of a source without an eigenfrequency (a charge or a multipole), radiation in the medium, i.e., V–Ch radiation, occurs only at a supraluminal velocity defined by condition (1). However, in this case, the medium is assumed to be homogeneous and time-invariant. If the medium is inhomogeneous and/or varies with time, some radiation is also possible from a source uniformly moving with a subluminal velocity. This radiation, the possibility of which was first considered in 1945 [4], is called transition radiation.

The simplest case of transition radiation is as follows: a charge uniformly moving along a straight line with any velocity crosses a boundary between two media. The point of intersection of the charge trajectory with the boundary becomes a source of transition radiation. This conclusion is most evident in the situation where the charge is incident from the vacuum on a metal surface (with a high conductivity), which plays the role of a perfect mirror (Fig. 4). From electrodynamics, it follows that, under such conditions, the field of the charge in vacuum is a sum of the fields of the charge q moving in the vacuum in the absence of the mirror and a charge $-q$ moving in the mirror toward the charge q (i.e., with the velocity $-v$). The charge $-q$ is called the "image" of the charge q. When charge q crosses the metal boundary, it falls into a conducting medium and ceases to produce a field in the vacuum; the image $-q$ also disappears. Thus, from the viewpoint of an observer in the vacuum, the annihilation of the pair of charges q and $-q$ occurs at the instant of crossing the boundary. From the same electrodynamics, it is known that annihilation, as well as any acceleration of charges (in the case under study, both charges q and $-q$ are abruptly stopped at the boundary), should be accompanied by a radiation. This is the transition radiation for the case under consideration.

For a perfect mirror, the energy radiated into the vacuum is expressed as

$$W_1(\omega, \theta) = \frac{q^2 v^2 \sin^2\theta}{\pi^2 c^3 [1 - (v^2/c^2)\cos^2\theta]^2},$$

$$W_1(\omega) = 2\pi \int W_1(\omega, \theta) \sin\theta d\theta \qquad (19)$$

$$= \frac{q^2}{\pi c}\left[\frac{1 + v^2/c^2}{2v/c}\ln\left(\frac{1 + v/c}{1 - v/c}\right) - 1\right].$$

In the ultrarelativistic limit (as $v \longrightarrow c$), we have

$$W_1(\omega) = \frac{q^2}{\pi c}\ln\frac{2}{1 - v/c} = \frac{2q^2}{\pi c}\ln\frac{2E}{mc^2}, \qquad (20)$$

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \gg mc^2.$$

Formulas (19) and (20) are derived rather easily [5–7, 42]. However, in the general case of two media characterized by complex permittivities ε_1 and ε_2, the calculations are cumbersome [4, 5, 42], and I do not present here even their results. It should only be noted that the aforementioned "backward" transition radiation (see Fig. 4) is of no practical interest. Presumably, it may account for the optical glow of anticathodes of X-ray tubes. In principle, this transition radiation may be used for measuring the particle energy E, because this quantity is involved in Eq. (20) for the emitted energy. However, in Eq. (20), the dependence on E is logarithmic while the absolute value of the energy W_1 is fairly small. It was found (in 1959; [43, 44]) that, for ultrarelativistic particles, it is expedient to consider the "forward" transition radiation, i.e., the radiation in the direction of the particle velocity, for example, when the particle passes from a substance into the vacuum. In this case, high-frequency radiation should be emitted as well, and the total radiation energy of a particle with a charge q and mass m is

$$W_2 = \int W_2(\omega) d\omega = \frac{q^2 \omega_p}{3c}\frac{E}{mc^2}, \qquad (21)$$

where ω_p is the plasma frequency of the substance (at high frequencies, all substances are equivalent to a plasma with a permittivity

$$\varepsilon = n^2 = 1 - \frac{\omega_p^2}{\omega^2}; \quad \omega_p^2 = \frac{4\pi e^2 N}{m_e},$$

where N is the electron concentration in the substance and e and m_e are the charge and mass of an electron).

The radiation energy W_2 is proportional to the particle energy E. Hence, by measuring W_2 one can determine E, which is important for high-energy particle

physics. Here, it should be noted that the use of the V–Ch effect for the energy measurements at high energies is ineffective. The point is that, in the ultrarelativistic region where $v \longrightarrow c$, the Cherenkov angle θ_0 (see Eq. (2)) and the radiation intensity (Eq. (3)) are almost insensitive to the particle energy $E = mc^2/\sqrt{1 - v^2/c^2}$. The measurement of the energy of the "forward" transition radiation W_2 is used as a basis for so-called transition counters, which have found wide application in high-energy particle physics [45, 46]. To avoid any misunderstanding, it should be noted that, since for a single boundary the energy W_2 (Eq. (21)) is rather small, the transition counters should use a "sandwich" of many sheets (plates) of a material with, e.g., air-filled gaps between them. The presence of many boundaries imposes certain limitations on the structure of such a counter. This fact involves some rather interesting physics (the consideration of the zone of the radiation formation) (see [5–7, 14, 42]).

6. TRANSITION RADIATION (THE GENERAL CASE). TRANSITION SCATTERING. TRANSITION BREMSSTRAHLUNG

The transition radiation that occurs at the intersection of a sharp interface represents the simplest case. In the general case, the transition radiation always appears when a source (charge) uniformly moves in an inhomogeneous and/or nonstationary medium or near it. In addition to the aforementioned situation with the "annihilation" of the source and its image, the transition radiation can be interpreted in a more general way. As an example, let us consider an isotropic transparent medium characterized by a refraction index n. Then, in the general case, the phase velocity of light in the medium is $v_{ph} = c/n(\omega, \mathbf{r}, t)$, where $\mathbf{r}$ represents the coordinates and t is time (evidently, in a homogeneous stationary medium, we have $n(\omega, \mathbf{r}, t) = n(\omega)$). The light radiation from a charge moving with a velocity v is governed by the ratio $v/v_{ph} = vn/c$. In vacuum, $n = 1$ and, at $v = \text{const}$, the radiation is absent (we assume that $v < c$); the radiation is possible only with the acceleration of the charge, i.e., when $v = v(t)$ and the acceleration is $w = dv/dt \neq 0$. In the medium, in the case of a uniform rectilinear motion with $v = \text{const}$ and $w = 0$, the ratio vn/c still may change because of the dependence of n on $\mathbf{r}$ and/or on t. Precisely this is the transition radiation, provided that the refraction index $n(\omega, \mathbf{r}, t)$ varies at the charge site or near it (within the zone of the radiation formation).

In the case of crossing a boundary between two media, the index n changes at this boundary. Another version involves any inhomogeneous medium (emulsion, plasma in an inhomogeneous magnetic field, etc.). One more possibility is of interest but not really practical: a charge uniformly moves in a homogeneous medium but, at some instant $t = t_0$ (or within some interval of time near the instant t_0), the refraction index changes in the whole medium, for example, because of the compression of the medium. Then, the point where the charge occurs at the instant t_0 plays (although not literally) a role similar to that of a boundary between two media [47, 42]. An important case of an inhomogeneous medium is a periodically inhomogeneous one, for example, a stack of plates used in transition counters [48, 42]. The transition radiation arising under such conditions is sometimes called resonance transition radiation or transition scattering. When a charge moves through a periodically inhomogeneous medium (a sine medium (see Eq. (22) below), a medium consisting of a set of sharp boundaries, etc.), one can say (from the charge standpoint) that a permittivity (refraction index) wave is incident on the charge. The scattering of this wave from the charge gives rise to transition radiation. However, the use of the term "transition scattering" would be not justified without the presence of this effect for a charge at rest. In this case, the term "transition radiation" seems to be inappropriate, while the term "transition scattering" fits the situation. For example, the effect occurs when a permittivity wave is incident on a charge q at rest (a fixed charge), and, as a result, an electromagnetic wave propagates (is scattered) from the charge (Fig. 5).

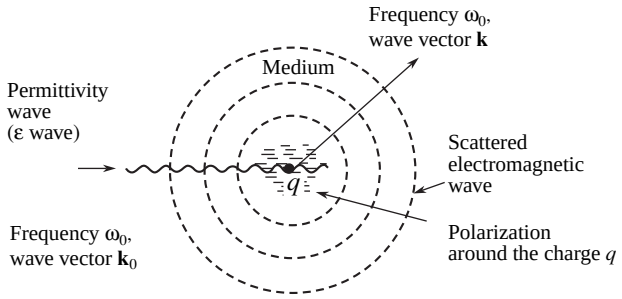


Fig. 5. Schematic diagram of the transition scattering of a permittivity wave from a stationary (fixed) charge q.

This result can also be easily understood without the general theory of transition radiation. For example, let us consider an isotropic transparent medium with a permittivity $\varepsilon = n^2$. If an acoustic wave propagates in such a medium, the density of the latter has the form $\rho = \rho^{(0)} + \rho^{(1)}\sin(\mathbf{k}_0\mathbf{r} - \omega_0 t)$, where $\mathbf{k}_0$ and ω_0 are the wave vector and the frequency of the acoustic wave, respectively. The variation of the density ρ is accompanied by a variation of ε, which gives rise to a permittivity wave:

$$\varepsilon(\mathbf{r}, t) = \varepsilon^{(0)} + \varepsilon^{(1)}\sin(\mathbf{k}_0\mathbf{r} - \omega_0 t), \qquad (22)$$

where $\varepsilon^{(0)}$ is the permittivity in the absence of the acoustic wave and $\varepsilon^{(1)}$ is the variation in ε due to the density variation. Evidently, a permittivity wave may be caused not only by an acoustic wave but also by some other factor, e.g., a longitudinal plasma wave.

Let us place a fixed or an infinitely heavy charge q in the medium. An electric field $\mathbf{E}$ and an induction $\mathbf{D} =$

$\varepsilon\mathbf{E}$ appear around the charge. If no wave is present, the field $\mathbf{E}$ is a Coulomb one:

$$\mathbf{E}^{(0)} = \frac{q\mathbf{r}}{\varepsilon^{(0)} r^3}, \quad D^{(0)} = \varepsilon^{(0)}\mathbf{E} = \frac{q\mathbf{r}}{r^3}. \tag{23}$$

In the presence of wave (22), in the first approximation (under the condition $|\varepsilon^{(1)}| \ll \varepsilon^{(0)}$), an additional polarization arises:

$$\delta\mathbf{P} = \frac{\delta\mathbf{D}}{4\pi} = \frac{\varepsilon^{(1)}}{4\pi}\mathbf{E}^{(0)}\sin(\mathbf{k}_0\mathbf{r} - \omega_0 t). \tag{24}$$

Such a polarization, which has no spherical symmetry (at $k_0 \neq 0$), gives rise to an electromagnetic wave of frequency ω_0 diverging from the charge (see Fig. 5). The wave number of this wave is $k = 2\pi/\lambda = (\omega_0/c)\sqrt{\varepsilon^{(0)}}$. If the permittivity wave is caused by an acoustic wave (as we assumed above), we have $k \ll k_0 = \omega_0/u$, where u is the velocity of sound (evidently, we assume that $u \ll c/\sqrt{\varepsilon^{(0)}}$).

The arising electromagnetic wave can be interpreted as a scattered one in the sense common to other types of scattering, for example, the Thomson scattering of an electromagnetic wave from an electron at rest (in this case, at rest means without taking into account the effect of the incident wave). If the medium is an isotropic plasma and the incident wave is a longitudinal (plasma) wave, the transition scattering under discussion represents a transformation of a longitudinal wave into an electromagnetic (transverse) one. This suggests that the transition scattering plays an important part in plasma physics, which actually is true [5–7, 42]. Let us illustrate this by an example. A longitudinal wave in plasma (its frequency is close to $\omega_p = \sqrt{4\pi e^3 N/m_e}$) contains some electric field and also involves a variation in ε. Thus, when a longitudinal wave propagates in plasma, the plasma particles (electrons and ions) are under the simultaneous effects of the electric field wave and the permittivity wave. Electrons of the plasma oscillate in the electric field and, hence, are sources of scattered electromagnetic waves (the so-called Thomson scattering), whose intensity is inversely proportional to the square of the mass of the scattering particle m. Therefore, the Thomson scattering from ions is characterized by an intensity that is $(m_i/m_e)^2$ times smaller than the intensity of the scattering from electrons (m_e is the electron mass and m_i is the ion mass). Hence, even in the case of Thomson scattering from the lightest ions, namely, protons with a mass $m_p = 1836m_e$, its intensity is $(1836)^2 \approx 3.4 \times 10^6$ times smaller than that in the case of scattering from electrons. By contrast, transition scattering in the first approximation does not depend on the mass of the scattering particle m and is also present at $m \longrightarrow \infty$. Therefore, in plasma, the total scattering of the longitudinal wave from ions actually is a transition scattering, whose intensity is of the same order of magnitude as the intensity of the longitudinal wave scattering from electrons. On the whole, without taking into account the transition scattering, an analysis of the processes that occur in plasma is impossible.

Another effect related to transition scattering is transition bremsstrahlung [50, 42]. Conventional bremsstrahlung occurs as follows: when particles collide, they are accelerated (or decelerated) and, as a result, emit electromagnetic waves. Since light particles (electrons) are accelerated more strongly than heavy particles (say, at the same velocity), the bremsstrahlung of electrons is much more intense (under comparable conditions) than the bremsstrahlung of heavy particles (protons, etc.). However, the aforesaid is valid only when the collisions and the corresponding bremsstrahlung occur in a vacuum. In the presence of a medium, the situation is different. As was mentioned above, radiation (transition radiation) is possible without any acceleration of the particles. Therefore, if a charge q moves in a medium (plasma) past a charge q', radiation appears even without any noticeable acceleration of some of these charges. This radiation is called transition bremsstrahlung. The physical nature of transition bremsstrahlung can be most easily understood by decomposing the field $\mathbf{E}$ and polarization $\mathbf{P} = [(\varepsilon - 1)/4\pi]\mathbf{E}$ of a uniformly moving charge q into waves with a wave vector $\mathbf{k}_0$; the frequency of these waves is $\omega_0 = \mathbf{k}_0\mathbf{v}$, where $\mathbf{v}$ is the velocity of the charge. These waves give rise to permittivity waves with the same ω_0 and $\mathbf{k}_0$. The permittivity waves are scattered from the other charges q', which results in transition bremsstrahlung.

Transition radiation, as well as transition scattering and bremsstrahlung, which are closely related to it, are considered in many publications and in a special book [42].

In this lecture, the problems were only briefly reviewed. However, I hope that it has clearly demonstrated the significance of the given area of research for physics (in the case of transition radiation, the transition counters and the applications in plasma physics are of special importance).

7. CONCLUDING REMARKS

In the development of physics and (evidently) other sciences, analogies, i.e., the transfer of notions from one area to another, play an important part. Therefore, for fruitful work in science it is important to have a wide scope of scientific interests rather than to only specialize (which often happens) in some narrow field of research. This rather trivial statement was formulated in my book [20] and, as I dare to believe, has been implemented in my scientific activities. The circle of problems described in this lecture may serve as an illustration of the above statement. Namely, the V–Ch effect is an analog of the Mach supersonic radiation (cone), the excitation of mechanical vibrations in supersonic

flows is analogous to the anomalous Doppler effect, and different types of transition radiation are also connected by common notions. On the whole, one can say that the analysis of different problems and effects related to radiation from uniformly moving sources forms a certain "ideology" and has its own "language." This can be seen from a number of examples, some of which are given above and others of which are given below (see also [5–7, 14, 42]; a clear popular description is given in [51]).

In 1946, Landau found that, in an isotropic plasma, even in the absence of collisions some attenuation of longitudinal (plasma) waves takes place [52]. This effect, which is called "Landau damping" (or collisionless damping), plays an important role in the physics of plasma and plasma like media (specifically, in the physics of metals and semiconductors, i.e., materials in which the conduction electrons form a kind of plasma). Landau obtained his result without any recourse to V–Ch radiation, and, undoubtedly, the mechanism of Landau damping can be understood without any references to the V–Ch effect. At the same time, the condition of Landau damping is the V–Ch condition (Eq. (8)) for the emission of a longitudinal wave by an electron (in this case, n in Eq. (8) is the refraction index for the longitudinal wave). Thus, for those who understand the mechanism of V–Ch radiation, the nature of Landau damping is evident.

Above, I have stressed that V–Ch radiation and the Doppler effect can be observed not only in the case of the motion of sources through a medium but also in the case of their motion in a narrow empty channel passing through the medium or near the medium boundary. The same is true for transition radiation and transition scattering. For example, let a charge uniformly move along a straight line over a flat surface of a medium consisting of two different materials. Then, when the charge passes over the boundary between these two media, transition radiation arises. On the whole, it always appears when some inhomogeneities occur near the trajectory of the charge, for example, when the charge enters or exits a metal waveguide (the inhomogeneity is represented by the edge of the waveguide), when the charge moves over a diffraction grating [53, 54], etc. This type of transition radiation is sometimes called diffraction radiation. The physical nature of this radiation can be most easily understood by using the aforementioned notion of the "image" charges, which move in the medium surrounding the charge trajectory (the "mirror"). The "images" move nonuniformly and emit radiation (another illustrative explanation of the effect is also possible; see, e.g., [51]).

As early as 60 years ago, at the first stage of the development of quantum electrodynamics, it became clear that, with allowance for quantum effects (primarily, electron–positron pair production, e^+e^-), the vacuum in a sufficiently strong magnetic field ceases to be the "absolute emptiness" of classical physics, in which electromagnetic waves of any frequency can freely propagate (without interacting with each other). By contrast, with allowance for the possibility of virtual pair production, the vacuum in a strong field behaves as a nonlinear anisotropic medium. The field is considered to be strong if it (e.g., the magnetic field H) is comparable to some characteristic field

$$H_c = \frac{m_e^2 c^3}{e\hbar} = 4.4 \times 10^{13}\ \text{Oe}. \qquad (25)$$

The characteristic electric field E_c is determined by the same Eq. (25), and its meaning is evident: within the Compton electron wavelength $\hbar/(m_e c) = 3 \times 10^{-11}$ cm, the field $2E_c$ performs work on the electron charge e that is equal to $2\hbar E_c/(m_e c) = 2m_e c^2$ and is necessary for the production of an e^+e^- pair (its mass at rest is equal to $2m_e c^2 \sim 10^{-6}$ erg $\sim 10^6$ eV). The field given by Eq. (25) is so strong that, for years, a nonlinear polarization of the vacuum seemed to be an abstract concept. However, in 1967 and 1968, magnetized neutron stars (pulsars) with typical fields of 10^{12}–10^{13} Oe were discovered. It was also found that, in semiconductors, it is possible to model to some extent the situation typical of strong fields (25) in a vacuum. Thus, strong fields have become an object available for astrophysical and physical studies. In the framework of this lecture, the aforesaid is of interest in relation to the fact that, in strong fields in vacuum, the V–Ch effect may occur, as may transition radiation and transition scattering (see [42] and the references given there). The vacuum also behaves as a medium in a gravitational field, which makes it possible to consider, e.g., transition scattering with a transformation from gravitational waves to electromagnetic ones [42].

In addition to the aforementioned acoustic analog of the V–Ch effect, acoustic analogs also exist for the electromagnetic transition radiation and transition scattering [55]. For me, a somewhat unexpected feature proves to be the important role that is played by the transition elastic-wave radiation in elastic systems, for example, in the case of the interaction between an inhomogeneous railway track and the wheels of a uniformly moving car [56].

Presumably, analogs of the V–Ch and Doppler effects and of transition radiation and scattering are possible for wave fields of any type and, hence, with allowance for the quantum theory, for particles of any type with a transformation (radiation) of fields (particles) of some other type. An example is the transition radiation in the form of electron–positron pair production that arises when a charge crosses some boundary, e.g., the boundary of an atomic nucleus. In brief, the radiation that accompanies a uniform motion of different kinds of sources is a universal phenomenon rather than an exotic effect. Therefore, it is natural that new experimental and theoretical studies concerned with this subject continue to appear in the literature. The papers I have seen in 1995 are as follows: the transition

(diffraction) radiation accompanying the motion of relativistic electrons over a diffraction grating [54], transition radiation in elastic systems [56], the transition radiation from a neutrino with a magnetic moment [57], the development of the theory of transition radiation [58, 59], the problem of the polarization of transition bremsstrahlung in plasma [60], and a detailed consideration of the transition scattering processes in an analysis of the bremsstrahlung in plasma [61] with an important application to the solar neutrino problem [62].

Thus, the area of physical research that appeared at the Lebedev Physical Institute of the Russian Academy of Sciences more than 50 years ago [1–4] and that was described in this lecture has now become an integral part of modern physics.

REFERENCES

1. P. A. Cherenkov, Dokl. Akad. Nauk SSSR **2**, 451 (1934).
2. S. I. Vavilov, Dokl. Akad. Nauk SSSR **2**, 457 (1934).
3. I. E. Tamm and I. M. Frank, Dokl. Akad. Nauk SSSR **14**, 107 (1937).
4. V. L. Ginzburg and I. M. Frank, Zh. Éksp. Teor. Fiz. **16**, 15 (1946); J. Phys. (Moscow) **9**, 353 (1945).
5. V. L. Ginzburg, *Theoretical Physics and Astrophysics*, 3rd ed. (Nauka, Moscow, 1987; Pergamon Press, Oxford, 1979).
6. V. L. Ginzburg, Tr. Fiz. Inst. im. P. N. Lebedeva, Akad. Nauk SSSR **176**, 3 (1986).
7. V. L. Ginzburg, Prog. Opt. **32**, 267 (1993).
8. O. Heaviside, Electrician, No. 23, 83 (1988); Philos. Mag. **27**, 324 (1889).
9. A. A. Tyapkin, Usp. Fiz. Nauk **112**, 731 (1974) [Sov. Phys. Usp. **17**, 288 (1974)].
10. T. R. Kaiser, Nature **247**, 400 (1974).
11. A. Sommerfeld, Gottinger Nachr. 99 (1904); Gottinger Nachr. 363 (1904); Gottinger Nachr. 201 (1905).
12. *Recollections on I. E. Tamm* (IzdAt, Moscow, 1995) [in Russian].
13. V. L. Ginzburg, Vestn. Ross. Akad. Nauk **65**, 848 (1995).
14. I. M. Frank, *The Vavilov–Cherenkov Radiation: Theoretical Problems* (Nauka, Moscow, 1988) [in Russian].
15. N. A. Dobrotin, E. L. Feĭnberg, and M. V. Fok, Priroda (Moscow), No. 11, 58 (1991).
16. P. A. Cherenkov, Dokl. Akad. Nauk SSSR **14**, 99 (1937); **14**, 103 (1937); **21**, 323 (1938).
17. I. E. Tamm, Usp. Fiz. Nauk **68**, 387 (1959); in *Collection of Scientific Works* (Nauka, Moscow, 1975), Vol. 1, p. 121 [in Russian].
18. B. M. Bolotowsky and Yu. N. Vavilov, Phys. Today **48** (12), 11 (1995).
19. P. A. Cherenkov, Phys. Rev. **52**, 378 (1937).
20. V. L. Ginzburg, *On Physics and Astrophysics* (Byuro Kvantum, Moscow, 1995) [in Russian].
21. E. L. Feĭnberg, Nauka i Zhizn', No. 8, 34 (1990).
22. V. P. Zrelov, *The Vavilov–Cherenkov Radiation and Its Application in High-Energy Physics* (Atomizdat, Moscow, 1968), Vols. 1 and 2 [in Russian].
23. *Cherenkov Detectors and Their Application in Science and Engineering* (Nauka, Moscow, 1990) [in Russian]; CERN Courier **34** (1), 22 (1994).
24. I. E. Tamm, J. Phys. (Moscow) **1**, 439 (1939); in *Collection of Scientific Works* (Nauka, Moscow, 1973), Vol. 1, p. 77 [in Russian].
25. V. L. Ginzbrug and I. M. Frank, Dokl. Akad. Nauk SSSR **56**, 699 (1947).
26. V. L. Ginzburg, Dokl. Akad. Nauk SSSR **56**, 145 (1947).
27. V. L. Ginzburg, Dokl. Akad. Nauk SSSR **56**, 253 (1947).
28. V. L. Ginzburg, Izv. Akad. Nauk SSSR, Ser. Fiz. **11**, 165 (1947).
29. E. Fermi, Phys. Rev. **57**, 485 (1940).
30. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media*, 3rd ed. (Nauka, Moscow, 1992; Pergamon, New York, 1984).
31. V. L. Ginzburg, Dokl. Akad. Nauk SSSR **24**, 130 (1939).
32. V. L. Ginzburg, Zh. Éksp. Teor. Fiz. **10**, 608 (1940); J. Phys. (Moscow) **3**, 101 (1940).
33. V. E. Pafomov, Tr. Fiz. Inst. im. P. N. Lebedeva, Akad. Nauk SSSR **16**, 94 (1961).
34. B. M. Bolotovskiĭ, Usp. Fiz. Nauk **62**, 201 (1957); Usp. Fiz. Nauk **75**, 295 (1961) [Sov. Phys. Usp. **4**, 781 (1962)].
35. D. A. Kirzhnits, *Some Problems in Nuclear Physics: in Memory of the 80th Anniversary of I. M. Frank* (Nauka, Moscow, 1989), p. 144 [in Russian].
36. V. L. Ginzburg, Zh. Éksp. Teor. Fiz. **10**, 589 (1940); J. Phys. (Moscow) **2**, 441 (1940).
37. I. M. Frank, Izv. Akad. Nauk SSSR, Ser. Fiz. **6**, 3 (1942).
38. V. L. Ginzburg and I. M. Frank, Dokl. Akad. Nauk SSSR **56**, 583 (1947).
39. V. L. Ginzburg and V. Ya. Éĭdman, Zh. Éksp. Teor. Fiz. **36**, 1823 (1959) [Sov. Phys. JETP **9**, 1300 (1959)].
40. B. E. Nemtsov, Zh. Éksp. Teor. Fiz. **91**, 44 (1986) [Sov. Phys. JETP **64**, 25 (1986)]; Izv. Vyssh. Uchebn. Zaved., Radiofiz. **28**, 1549 (1985); Izv. Vyssh. Uchebn. Zaved., Radiofiz. **30**, 968 (1987); B. E. Nemtsov and V. Ya. Eĭdman, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **27**, 1192 (1984); Izv. Vyssh. Uchebn. Zaved., Radiofiz. **30**, 226 (1987).
41. V. L. Ginzburg and V. P. Frolov, Pis'ma Zh. Éksp. Teor. Fiz. **43**, 265 (1986) [JETP Lett. **43**, 339 (1986)]; Phys. Lett. A **116**, 423 (1986); Tr. Fiz. Inst. im. P. N. Lebedeva, Akad. Nauk SSSR **197**, 8 (1989).
42. V. L. Ginzburg and V. N. Tsytovich, *Transition Radiation and Transition Scattering* (Nauka, Moscow, 1984; Hilger, Bristol, 1990).
43. G. M. Garibyan, Zh. Éksp. Teor. Fiz. **37**, 527 (1959) [Sov. Phys. JETP **10**, 372 (1960)]; G. M. Garibyan and Yan Shi, *X-ray Transition Radiation* (Akad. Nauk Arm. SSR, Yerevan, 1983).
44. K. A. Barsukov, Zh. Éksp. Teor. Fiz. **37**, 1106 (1959) [Sov. Phys. JETP **10**, 787 (1960)].
45. C. W. Fabjan and H. G. Fischer, Rep. Prog. Phys. **43**, 1003 (1980).
46. K. Kleinkhecht, Phys. Rep. **84**, 85 (1982).
47. V. L. Ginzburg, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **16**, 512 (1973).

48. M. L. Ter-Mikaelyan, *Influence of the Medium on Electromagnetic Processes at High Energies* (Akad. Nauk Arm. SSR, Yerevan, 1969) [in Russian]; R. A. Bagiyan and M. L. Ter-Mikaelyan, Zh. Éksp. Teor. Fiz. **81**, 1249 (1981) [Sov. Phys. JETP **54**, 666 (1981)].
49. V. L. Ginzburg and V. N. Tsytovich, Zh. Éksp. Teor. Fiz. **65**, 1818 (1973) [Sov. Phys. JETP **38**, 909 (1973)].
50. V. N. Tsytovich, Tr. Fiz. Inst. im. P. N. Lebedeva, Akad. Nauk SSSR **66**, 176 (1973).
51. B. M. Bolotovskiĭ and V. A. Davydov, *Charge, Medium, and Radiation* (Znanie, Moscow, 1989) [in Russian].
52. L. D. Landau, Zh. Éksp. Teor. Fiz. **16**, 574 (1946).
53. S. J. Smith and E. M. Purcell, Phys. Rev. **92**, 1069 (1953).
54. K. J. Woods, J. E. Walsh, R. E. Stoner, *et al.*, Phys. Rev. Lett. **74**, 3808 (1995).
55. V. I. Pavlov and A. I. Sukhorukov, Usp. Fiz. Nauk **147**, 83 (1985) [Sov. Phys. Usp. **28**, 784 (1985)].
56. A. I. Vesnitskiĭ, A. V. Kononov, and A. V. Metrikin, Prikl. Mekh. Tekh. Fiz. **36**, 170 (1995); A. I. Vesnitskiĭ and A. V. Metrikin, Usp. Fiz. Nauk **166**, 1043 (1996) [Phys. Usp. **39**, 983 (1996)].
57. M. Sakuda and Y. Kurihara, Phys. Rev. Lett. **74**, 1284 (1995).
58. V. V. Krechetov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **38**, 639 (1995).
59. I. I. Kalikinskiĭ, Zh. Tekh. Fiz. **65** (10), 131 (1995) [Tech. Phys. **40**, 1047 (1995)].
60. V. B. Korsakov and G. D. Fleĭshman, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **38**, 887 (1995).
61. V. N. Tsytovich, Usp. Fiz. Nauk **165**, 89 (1995) [Phys. Usp. **38**, 87 (1995)].
62. V. N. Tsytovich *et al.* (for the Central Laboratory of the Research Councils, England), Technical Report RAL-TR 95-066; V. N. Tsytovich, R. Bingham, U. De Anglls, and A. Forlani, Usp. Fiz. Nauk **166**, 113 (1996) [Phys. Usp. **39**, 103 (1996)].

Translated by E. Golyamina

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 110–112. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 133–134.

SHORT COMMUNICATIONS

The Potential Component of the Field Generated by an Instability of a Shear Flow

S. A. Rybak

Andreev Acoustics Institute, Russian Academy of Sciences, ul. Shvernika 4, Moscow, 117036 Russia

e-mail: rybak@akin.ru

Received May 24, 2004

The set of hydrodynamic equations for a two-dimensional model with (x, y) coordinates, where x is the coordinate along the flow $U(y)$ and y is the coordinate along the flow normal, has the following form in the quadratic approximation [1]:

$$\frac{\partial}{\partial t}v_i + \left[(v_k + U\delta_{kx})\frac{\partial}{\partial x_k}(v_i + U\delta_{ix})\right]$$

$$= -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \eta\Delta v_i + \left(\varsigma + \frac{\eta}{3}\right)\frac{\partial}{\partial x_i}\frac{\partial v_k}{\partial x_k};$$

$$\frac{\partial \rho}{\partial t} + U\frac{\partial \rho}{\partial x} + \frac{\partial}{\partial x_k}(\rho v_k) = 0; \qquad (1)$$

$$\delta\rho = \delta p/c^2 - \frac{\gamma - 1}{2\rho_0 c^4}\delta p^2;$$

$$v_x = \frac{\partial \phi}{\partial x} + \frac{\partial \varphi}{\partial y}; \quad v_y = \frac{\partial \varphi}{\partial y} - \frac{\partial \phi}{\partial x}.$$

In the linearized limit, the set of equations has the form

$$D\Delta\phi - U_{yy}\frac{\partial \phi}{\partial x} + \Delta\varphi U_y + U_{yy}\frac{\partial \phi}{\partial y} = \eta\Delta^2\phi;$$

$$D\Delta\varphi + \frac{1}{\rho_0}\Delta p = \left(\varsigma + \frac{4}{3}\eta\right)\Delta^2\varphi - 2U_y\frac{\partial}{\partial x}\left(\frac{\partial \varphi}{\partial y} - \frac{\partial \phi}{\partial x}\right); \qquad (2)$$

$$\frac{1}{\rho c^2}Dp + \Delta\varphi = 0; \quad D = \frac{\partial}{\partial t} + U\frac{\partial}{\partial x}.$$

The first equation of set (2) is a modified Orr–Sommerfeld equation, which, when taken without the two last terms on the left-hand side, is rather well-known. The equation written above takes into account the interaction that occurs between the eddy and potential components of the velocity field because of the presence of the shear flow $U(y)$. The second and third equations taken together are equivalent to the Lighthill equation [3] with the linearized right-hand side. They are written with allowance for a Doppler frequency shift that is nonuniform along the y axis. The instability appears already in the case of a linear variation of the flow in the y direction:

$$U(y) = U_0 + \varepsilon y, \quad U_{yy}(y) = 0. \qquad (3)$$

Below (Eq. (5)), it will be shown that the inclusion of the potential component of the wave field is a necessary condition for the appearance of an instability in the absence of an inflection point in the flow profile.

Ignoring the viscosity, we obtain

$$D\Delta\phi + \Delta\varphi U_y = 0;$$

$$D\Delta\varphi + \frac{1}{\rho_0}\Delta p = -2U_y\frac{\partial}{\partial x}\left(\frac{\partial \varphi}{\partial y} - \frac{\partial \phi}{\partial x}\right); \qquad (4)$$

$$\frac{1}{\rho_0 c^2}Dp + \Delta\varphi = 0.$$

Equations (4) yield a closed equation for $p(x, y, t)$:

$$\Delta\left(\Delta p - \frac{1}{c^2}D^2 p\right) = 2\frac{U_y}{c^2}\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}Dp - U_y\frac{\partial}{\partial x}p\right). \qquad (5)$$

The presence of derivatives of different parity in Eq. (5) is a feature of a nonconservative problem.

Note that Eqs. (4) yield a Poisson equation for the stream function:

$$\Delta\phi = \frac{U_y}{\rho c^2}p. \qquad (6)$$

According to Eqs. (3), we obtain

$$U_y = \varepsilon.$$

Solving set of equations (4) to the first order in ε, i.e., assuming that U and U_y are constant, we obtain

$$\varphi = c_1\exp[j(-\omega t + kx + qy)],$$
$$\phi = c_2\exp[j(-\omega t + kx + qy)]. \qquad (7)$$

1063-7710/05/5101-0110$26.00

Assume that

$$\omega = \omega_0 + \varepsilon\omega_1. \tag{8}$$

From Eqs. (4), we obtain a dispersion equation:

$$\begin{aligned}&(\omega - Uk^2) - (k^2 + q^2)c^2 \\ &= \frac{-2j(\omega - Uk)U_y kq + 2k^2U_y^2}{k^2 + q^2}.\end{aligned} \tag{9}$$

Equation (9) yields

$$\begin{aligned}\omega_0 &= U_0 k + c(k^2 + q^2)^{1/2}, \\ \omega_1 &= j\frac{U_y kq}{k^2 + q^2}.\end{aligned} \tag{10}$$

Here, the last term in Eq. (9), i.e., the term quadratic in U_y, is ignored. Note that, when

$$c^2 \longrightarrow \infty,$$

we have

$$k^2 \longrightarrow -q^2,$$

and solution (10) fails.

Now, let us return to set of equations (1) and linearize it without separating the velocity field into the potential and eddy components. Then, we obtain a wave equation for the quasi-harmonic pressure field $p(x, y, t)$ in a form simpler than that of Eq. (5) even for an arbitrary $U(y)$:

$$\begin{aligned}&\left(\frac{\omega_d^2}{c^2} + \Delta\right)p = -\frac{2kU'}{\omega_d}p_y, \\ &p = p(y)\exp(-j\omega_d t + jkx), \\ &\omega_d = \omega - U(y)k.\end{aligned} \tag{11}$$

It should be stressed that Eq. (5) describes the potential component of a quasi-eddy wave, while Eq. (11) describes the potential component of a quasi-potential wave. In both cases, the instability of the shear flow generates both potential and eddy components. Hence, linearized set of equations (2) corresponds to four waves: an eddy wave accompanied by a potential wave attached to it and a potential wave accompanied by an eddy wave. Their relation is described by the first equation of set (2).

In the linearized limit, we have

$$\omega_d = \omega_{d0} - U'yk, \quad \omega_{d0} = \omega - U_0 k.$$

Here,

$$p_y = \frac{\partial p}{\partial y} = -j\rho\omega_d v_y.$$

Let us determine $p(y)$ for a normal wave in a waveguide given as $0 < y < h$ to first order in U' ($U_0 = 0$):

$$\begin{aligned}p(y) &= A\exp(jq_1 y) + B\exp(-jq_2 y), \\ q_{1,2} &= -\frac{jkU'}{\omega_{d0}}(+/-)\left(\frac{\omega_{d0}^2}{c^2} - k^2\right)^{1/2}.\end{aligned} \tag{12}$$

In fact, we use two small parameters:

$$\frac{U'hk}{\omega_{d0}} \ll 1, \quad \frac{h\omega_{d0}^2}{qc^2} \ll 1.$$

Let us determine the solution under the boundary conditions

$$\begin{aligned}v_y &= 0 \text{ at } y = 0, \\ p &= Z(\omega)v_y \text{ at } y = h.\end{aligned} \tag{13}$$

The dispersion equation takes the form

$$\begin{aligned}&\exp(2jq_0 h)\frac{[1 - Z_1(q_0 - P)](q_0 + P)}{[1 + Z_1(q_0 + P)](q_0 - P)} + 1 = 0; \\ &P = \frac{jkU'}{\omega_{d0}}, \quad q_0 = \left(\frac{\omega_{d0}^2}{c^2} - k^2\right)^{1/2}; \\ &Z_1 = \frac{Z}{-j\omega_d\rho}.\end{aligned} \tag{14}$$

Solving Eq. (14), we obtain

$$\mathrm{Im}\,\omega(q_0) < 0. \tag{15}$$

The following specific cases are possible for the impedance at the boundary:

(a) $\mathrm{Im}Z \neq 0$

and

(b) when $Z = \infty$, $\exp(2jq_0 h) = 1$, $q_{0n} = \dfrac{\pi n}{h}$.

In the first case, four solutions ω_{d0} to Eq. (12) are possible, and one of them has an imaginary part satisfying condition (15).

Earlier [4] it was shown that, for linearized set of equations (1) in the presence of a shear flow $U(y)$, the energy integral has the form

$$\begin{aligned}&\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\left[\frac{1}{2}\rho(v_x^2 + v_y^2) + \frac{p^2}{2\rho c^2}\right] \\ &+ \frac{\partial}{\partial x_i}(pv_i) = -\rho v_x v_y U_y.\end{aligned} \tag{16}$$

The energy flux from the shear flow, which is the origin of the instability, is described in terms of the Reynolds stresses on the right-hand side of Eq. (16). Its sign is determined by the sign of the right-hand side of

Eq. (16). Both components of the unstable wave, (ψ, φ), are the consequence of this process. No change of sign is necessary for the curvature of the flow profile.

At the same time, a change of sign of the curvature of the flow profile is necessary for the appearance of an instability if the potential component of the instability-generated field is not taken into account.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 02-02-17143, and the Program for Supporting the Leading Scientific Schools, grant no. NSh-1176-2003-2.

REFERENCES

1. S. A. Rybak, Akust. Zh. **48**, 854 (2002) [Acoust. Phys. **48**, 756 (2002)].
2. C.-C. Lin, *The Theory of Hydrodynamic Stability*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1966; Inostrannaya Literatura, Moscow, 1958).
3. J. Lighthill, *Waves in Fluids* (Cambridge Univ. Press, Cambridge, 1978; Mir, Moscow, 1981).
4. S. A. Rybak, in *Proceedings of 16th ISNA* (Moscow, 2002), Vol. 1, p. 235.

Translated by E. Golyamina

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 24–32. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 37–46.

Akusticheskii Zhurnal and the Andreev Acoustics Institute

N. A. Dubrovsky and I. B. Esipov

Andreev Acoustics Institute, Russian Academy of Sciences, ul. Shvernika 4, Moscow, 117036 Russia

e-mail: dubrov@akin.ru; ibesipov@akin.ru

Received June 23, 2004

Abstract—In 2003, the Andreev Acoustics Institute celebrated its 50th birthday. The formation of the institute in the framework of the Academy of Sciences of the USSR, which occurred in 1953, lead to the appearance of the *Akusticheskii Zhurnal* (*Acoustical Physics*). This journal always reflected the main achievements of the scientists from the Acoustics Institute. The leading scientists of the institute, in their turn, always took an active part in the work of the editorial board of the journal. In the present paper, the history of the Acoustics Institute is briefly reviewed and its current areas of research are outlined. The role of the institute in the development of acoustical research in Russia is discussed. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

The establishment of the *Akusticheskii Zhurnal* in 1955 was a natural consequence of the foundation of the Acoustics Institute of the Academy of Sciences of the USSR in December 1953. Thus, the Acoustics Institute and the *Akusticheskii Zhurnal* (*Acoustical Physics*) almost simultaneously celebrate their 50th anniversaries. The need for a specialized academic journal devoted to acoustics was determined by the standard accepted at that time for the scientific activity of an academic institute. Each academic institute either had its own academic journal devoted to the corresponding area of research or published its own proceedings. A publication of scientific results in such a journal or in the proceedings of the institute was the final step of any research.

According to the decision of the Presidium of the Academy of Sciences of the USSR, the first editor-in-chief of the *Akusticheskii Zhurnal* was Academician N.N. Andreev, who initiated the foundation of the Acoustics Institute of the Academy of Sciences of the USSR. The first editorial board included scientists representing the acoustical elite of the country. They included the first director of the Acoustics Institute, Corresponding Member (at that time) of the Academy of Sciences of the USSR L.M. Brekhovskikh and the leaders of the main lines of research carried out at the institute V.S. Grigor'ev (deputy editor-in-chief) and L.D. Rozenberg who, together with the chair of the Department of Acoustics of Moscow State University S.N. Rzhevkin and Corresponding Member of the Academy of Sciences of the USSR S.Ya. Sokolov (representing the acousticians from Leningrad), determined the style of the journal and the requirements imposed on the quality of publications. A substantial contribution to the formation and development of the *Akusticheskii Zhurnal* was made by L.M. Lyamshev. Beginning in 1963, he was deputy editor-in-chief and, from 1987 to 2002, editor-in-chief of the journal. Actually, it is in this period of time that the current image of the journal was formed.

The very first issue of the journal began with an address from the editorial board to the readers, which said:

"In view of the rapid development of scientific and engineering acoustics in the USSR and the need for the open publication of the results of research and engineering works in acoustics, the Presidium of the Academy of Sciences of the USSR has acknowledged the necessity to begin in 1955 the publication of the *Akusticheskii Zhurnal* of the Academy of Sciences of the USSR.

The main purpose of the journal is to provide the information on research in theoretical and experimental acoustics and on the problems of the improvement and production of acoustical measuring equipment, the improvement of the quality of acoustical equipment and systems of sound transmission and reproduction, the improvement of the acoustic properties of public buildings and homes, noise and vibration control, and also on the problems of the practical and scientific applications of ultrasound…."

Research in these and other areas has been carried out at the Acoustics institute from the first days of its existence under the supervision of well-known specialists heading the main lines of the scientific activity of the institute.

The Acoustics Institute was founded according to the decision of the government of the USSR, which was issued on October 31, 1953. The institute was organized on the basis of the Acoustical Laboratory of the Lebedev Physical Institute of the Academy of Sciences

1063-7710/05/5101-0024$26.00 © 2005 Pleiades Publishing, Inc.

of the USSR. The laboratory, in its turn, was organized in 1939, also on the initiative of N.N. Andreev.

Thus, an indissoluble relation exists between the history of the development of the *Akusticheskii Zhurnal* and the history of the Acoustics Institute. Therefore, the role and place of the *Akusticheskii Zhurnal* in the development of acoustics in Russia cannot be completely understood without considering the history of the development of this institute.

SOME HISTORY

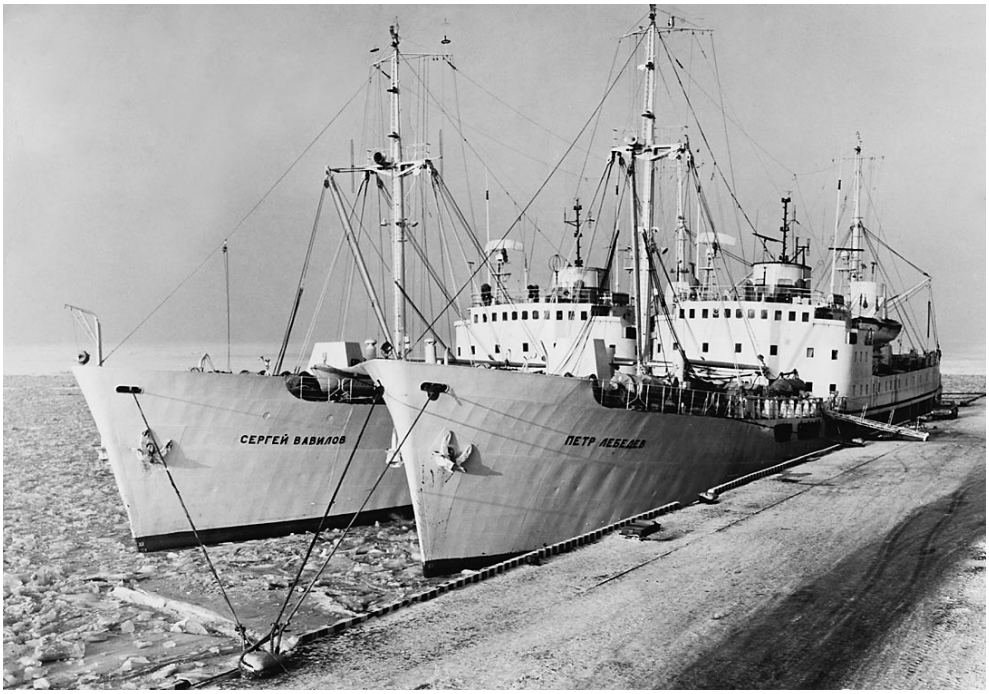
The most important basic result of research carried out by the Acoustical Laboratory of the Lebedev Physical Institute in the late 1940s and early 1950s was the discovery of long-range sound propagation in the ocean. The role of L.M. Brekhovskikh and L.D. Rozenberg, researchers from the Acoustical Laboratory, in this discovery was acknowledged with the First-degree Stalin Award, which was given to them and their colleagues in 1951. The Acoustics Institute was founded with the aim to continue research in this area, and the main purpose of the institute was to carry out comprehensive studies in underwater acoustics. The results of these studies should serve as the basis for a considerable increase in the range of underwater acoustic surveys. Decisive contributions to the development of Russian acoustics and underwater acoustics and to the formation of the main lines of research at the Acoustics Institute were made by the following leading scientists: N.N. Andreev, L.M. Brekhovskikh, V.S. Grigor'ev, M.A. Isakovich, G.D. Malyuzhinets, A.V. Rimskii-Korsakov, L.D. Rozenberg, Yu.M. Sukharevskii, and B.D. Tartakovskii. During World War II, most of these scientists were involved in the development and practical application of antimine acoustic trawls, sound detectors, and listening sonars for the air raid system and also in theoretical and experimental research for naval purposes.

Over the half-century of its history, the Acoustics Institute has been a national research center for basic and applied studies in many fields of modern acoustics and underwater acoustics. Since 1994, the institute has been a State Research Center of the Russian Federation.

Three main periods can be distinguished in the history of the institute. In the first period (from 1954 to 1961), the institute belonged to the Academy of Sciences of the USSR. At that time, the first buildings of the institute were constructed and the following research stations were organized: the Sukhumi marine research station on the Black Sea and the Volga research station on the Ivan'kovskoe storage lake (now, it has become the Institute of Applied Acoustics). In addition, the first specialized research vessels (*Sergei Vavilov* and *Petr Lebedev*) intended for acoustic studies in the ocean were built and given to the Acoustics Institute. The Department of Acoustics was organized at the Moscow Institute of Physics and Technology. To the graduate students of this department, lectures and laboratory training were given at the Acoustics Institute. In this period, several tens of young scientists—graduates of the best universities of the USSR—came to the Acoustics Institute and formed the basis of its staff. The development of research included the study of the formation of an underwater sound channel and the zonal structure of the sound field that was necessary for designing long-range hydroacoustic systems (Yu.M. Sukharevskii, I.D. Ivanov, and A.L. Sosedova). At the same time, regular All-Union Acoustical conferences were reestablished. The first of them was held in September 1931 at the Physicotechnical Institute (Leningrad) and was initiated by Academician N.N. Andreev. The second conference was held in 1958 in Moscow. Later, the conferences were held on a regular basis up to 1991. The Acoustics Institute was involved in organizing all of these conferences. Thus, from the very beginning, the Acoustics Institute not only carried out research but also supervised research and field tests at its branches, developed original methods of acoustic studies on research vessels in the ocean (no such experience existed in the world at that time), educated young specialists, and contributed to the popularization of acoustical research over the whole country.

Chairman of the Supreme Soviet of the USSR Kliment Efremovich Voroshilov gives the Star of the Hero of Socialist Labor to Academician Nikolai Nikolaevich Andreev.

The second period began in 1961, when the institute (together with some other institutes of the Academy of Sciences) was transferred from the Academy of Sciences to an industrial ministry. In this period (which is the longest one), new buildings were constructed for the institute, the pilot production department was increased, and new lines of research were started. The



Research ships of the Acoustics Institute, *Sergei Vavilov* and *Petr Lebedev*, with which unique studies in the ocean were carried out for a quarter of a century.

departments, laboratories, and branches of the institute received new scientists and engineers: by 1989, the staff of the institute together with its branches included 2852 people. The possibilities for experiments were extended, the Northern and Pacific branches of the institute and the Black Sea laboratory were formed, new research vessels were built (*Akademik Nikolai Andreev* and *Akademik Boris Konstantinov*), and unique experimental test facilities were constructed (the anechoic and reverberation chambers and a vibration-insulated chamber). A powerful computer center equipped with the best computers available at that time was formed. Within this period, unique data on the acoustic properties of the ocean were accumulated; these data were obtained from 44 expeditions on research ships. Many scientists and engineers from the Acoustics Institute received high state and governmental awards for their achievements. The USSR government and the Central Committee of the Communist Party decided to give the official status of the Leading institute in acoustics and underwater acoustics in the country to the Acoustics Institute. The directors of the institute of that time were N.A. Grubnik (1963 to 1980) and F.I. Kryazhev (1980 to 1989).

The beginning of the third period (since 1991) coincides with the fall of the Soviet Union and the formation of a new economic system in Russia. During this period, since 1989, the director of the institute has been N.A. Dubrovsky.

Researchers working at the Acoustics Institute are the authors of many papers that appear in the *Akusticheskii Zhurnal* (*Acoustical Physics*). From these publications, one can get an idea of the scientific results obtained in this period at the institute. Even the very first papers submitted by the scientists from the Acoustics Institute to the *Akusticheskii Zhurnal* were characterized by a very high scientific and professional level. The first issue of the journal began with the paper by N.N. Andreev "On Some Second-Order Quantities in Acoustics." This paper triggered numerous studies in nonlinear acoustics in the USSR. Andreev not only initiated but also supervised many of these studies. Thirty years later, a group of scientists received the USSR State Award for a series of works in this field of acoustics.

The first issue of the *Akusticheskii Zhurnal* also contained a paper by L.M. Brekhovskikh and I.D. Ivanov, who considered a particular type of attenuation for waves propagating in layered inhomogeneous media. The results presented in this paper were included in the widely known monograph by L.M. Brekhovskikh *Waves in Layered Media*, which appeared in 1957. There, readers can also find a paper by Yu.P. Lysanov, which proposes a new method for solving the problems of scattering from a periodically inhomogeneous surface. The clarity of the solution obtained by Lysanov stimulated experimental studies in this area of research, and the method itself was further developed by other authors.

The publication by I.P. Golyamina devoted to the study of the relation between the electric impedance and the acoustic radiation of piezoelectric plates, shows

that the development of this area of research was given much attention at the institute. The same issue of the *Akusticheskii Zhurnal* contained the paper by L.A. Chernov "Correlation of the Phase and Amplitude Fluctuations in the Wave Propagation through a Medium with Random Inhomogeneities." The problem of wave propagation in randomly inhomogeneous media has become the object of investigation for many scientists, and the papers by Chernov are still the most cited ones in this area of research. A number of interesting papers on wave diffraction were published by G.D. Malyuzhenets. He presented the results of his studies in the framework of the general Sommerfeld integral theory for stationary diffraction problems in an arbitrary region and proposed new mathematical methods for the diffraction theory.

The theoretical studies carried out at the institute were characterized by a wide scope of interests. For example, the theory of highly viscous microinhomogeneous media, which was developed by M.A. Isakovich and I.A. Chaban, still attracts considerable attention from researchers. Yu.L. Gazaryan was the first to derive an exact solution to the problem of a multiple scattering of sound in a one-dimensional randomly inhomogeneous medium. The same author solved the problem on the propagation of nuclear-explosion-caused infrasound in the atmosphere. The papers by A.D. Lapin on waveguide insulation attracted considerable interest. Remarkable investigations into the wave scattering from fractal structures were carried out by L.M. Lyamshev, Yu.P. Lysanov, and I.A. Urusovskii.

In the USSR, the development of active compensation methods for vibration and sound fields was pioneered by B.D. Tartakovskii. He is the author of publications concerned with the use of statistical methods for evaluating random vibration and sound fields and the propagation of vibrations through inhomogeneous structures. It should be noted that the first research projects carried out by Tartakovskii were concerned with the theoretical study of sound reflection from the shell and the sound-absorbing layer of the dome that should cover the Large Hall of the Palace of Soviets. Under the supervision of Tartakovskii and with his participation, vibration-absorbing polymer materials were designed. These materials are currently in use for noise control in production goods and transport.

The beginning of computational methods in acoustics is associated with the publications by V.Yu. Zavadskii. The class of problems studied by Zavadskii was closely related to sound propagation in the ocean, and his first scientific results were devoted to the development of effective methods for studying the fields in acoustic waveguides. He was the first to propose the methods for the numerical solution of the Helmholtz equation in open regions and waveguides on the basis of the finite-difference approximation of the differential operators of this equation. Prior to his studies, the finite-difference method was used only for bounded regions. Zavadskii also succeeded in developing rapidly converging iteration schemes for open regions with nonlocal boundary conditions representing the difference analogs of the radiation conditions.

In the late 1940s, under the supervision of Yu.M. Sukharevskii, research projects concerned with the development of antisonar protection for submarines were carried out. The development of the theory of wave diffraction by shells (G.D. Malyuzhenets) and the development of original sound-absorbing coatings (under the supervision of V.V. Tyutekin) resulted in the design of noise-suppressing coatings for submarines of the third and fourth generations (which are still in use). The main designers of these coatings were all from the Acoustics Institute: A.E. Vovk, Yu.B. Upadyshev, and I.P. Zhukov.

Following the tradition founded by Academician N.N. Andreev when he worked at the Acoustical Laboratory of the Lebedev Physical Institute, theoretical and experimental studies in physical acoustics were carried out. Here, one should note the studies on the physics of ultrasound and in ultrasonics (L.D. Rozenberg), in biological and nonlinear acoustics (N.N. Andreev), and in musical acoustics (A.V. Rimskii-Korsakov). Many of the applied studies carried out at the institute within the last ten years originate from these physical studies. For example, the research and development that resulted in the design of an ultrasonic imaging system were started at the institute in the 1960s by Yu.B. Semennikov. The design of pneumoacoustic atomizers proposed by Yu.Ya. Borisov and widely used today originates from studies initiated in the 1980s at the Ultrasonics Department of the institute.

During the half-century-long history of the Acoustics Institute, its staff members have carried out numerous investigations at the highest scientific level. The authors of many of these works received state prizes and awards. The contribution of the scientists from the Acoustics Institute to the development of underwater sonars for the Navy was several times honored with state awards. The authors of the most significant achievements in this area of research are L.M. Brekhovskikh (the Lenin prize of 1970) and Yu.M. Sukharevskii and V.I. Mazepov (the USSR State Award, 1967). For underwater acoustic studies in the Arctic, V.S. Grigor'ev, N.A. Grubnik, F.I. Kryazhev, and N.A. Petrov received the USSR State Award in 1969. In the 1960s, I.A. Viktorov initiated research in acoustoelectronics at the Acoustic Institute. Later, in 1974, Viktorov and his colleagues from the Institute of Radio Electronics of the Academy of Sciences of the USSR received the USSR State Award. The State Award of 1976 was given to the authors of the monograph *Ocean Acoustics*, edited by L.M. Brekhovskikh (Nauka, Moscow, 1974). Its authors were scientists from the Acoustics Institute: N.S. Ageeva, I.B. Andreeva, L.M. Brekhovskikh, V.I. Volovov, Yu.Yu. Zhitkovskii, Yu.P. Lysanov, A.V. Furduev, S.D. Chuprov, and R.F. Shvachko.

Laureates of the USSR State Award, authors of the monograph *Ocean Acoustics* (from left to right): Yu.Yu. Zhitkovskii, I.B. Andreeva, S.A. Chuprov, L.M. Brekhovskikh, N.S. Ageeva, R.F. Shvachko, Yu.P. Lysanov, A.V. Furduev, and V.I. Volovov.

In 1984, the USSR State Award was given to the research group working under the supervision of N.A. Dubrovsky on bioacoustic studies. In the 1970s, under the supervision of L.M. Lyamshev and K.A. Naugol'nykh, research in nonlinear acoustics was started at the Acoustics Institute. In the 1980s, this research was complemented with pioneering investigations in a new area of research: laser acoustics. At the same time, aerohydrodynamic acoustic studies were carried out under the supervision of A.V. Rimskii-Korsakov. The latter studies resulted in the formation of a combined team of researchers from the Acoustics Institute, Moscow State University, the General Physics Institute of the Russian Academy of Sciences, Taganrog Radio Engineering University, and the Institute of Applied Physics of the Academy of Sciences of the USSR. For these studies, L.M. Lyamshev, K.A. Naugol'nykh, and A.V. Rimskii Korsakov together with a group of coauthors received the USSR State Award in 1985. As a continuation of the achievements of scientists from the Acoustics Institute in this direction of research, one can cite the Russian Federation State Award given in 1997 to S.A. Rybak and his colleagues from other institutions for a series of works on the dynamics of intense noise waves and nonlinear structures.

THE ACOUSTICS INSTITUTE TODAY

Let us now briefly describe the State Research Center of the Russian Federation, Andreev Acoustics Institute in its present state.

The Andreev Acoustics Institute of today is a scientific institution that carries out comprehensive research and engineering projects in the key areas of modern acoustics: underwater and ocean acoustics, geoacoustics, processing and discrimination of acoustic signals, nonlinear acoustics, photoacoustics, development of active and passive methods and means for noise and vibration control, thermoacoustics, bioacoustics, medical acoustics, and ultrasonic applications.

The institute has gained world wide recognition owing to well-known discoveries and to leading scientists including Academician L.M. Brekhovskikh (ocean acoustics), Professor G.D. Malyuzhenets (theory of wave diffraction), Professor L.D. Rozenberg (physics of ultrasound and ultrasonic technologies), Professor A.V. Rimskii-Korsakov (acoustics of gas jets), Professor L.M. Lyamshev (photoacoustics), and Professor N.A. Dubrovsky (bioacoustics). The scientists and engineers from the Acoustics Institute have, in the fifty years of its existence, laid the foundation for the development of underwater sonars for three generations of submarines, surface warships, and stationary hydroacoustic systems (Yu.M. Sukharevskii, V.S. Grigor'ev, S.G. Gershman, N.A. Dubrovsky, V.I. Mazepov, O.P. Galkin, V.V. Ol'shevskii, R.Yu. Popov, Yu.I. Tuzhilkin, V.M. Baronkin, S.I. Dvornikov, V.P. Tebyakin, and many others).

Today, the staff of the institute includes 25 professors and doctors of science and 75 candidates of science (most of whom are well known to the international scientific community) and more than 100 engineers, com-

puter programmers, and experimental physicists. The institute retains its Northern branch in Severomorsk, where research for the Russian Navy is carried out. The institute has postgraduate and doctoral departments. Scientists from the institute successfully participate with their projects in various competitions of the Russian Foundation for Basic Research and international scientific foundations (CRDF, INTAS, Copernicus).

The scientific activity of the institute is coordinated by the Deputy Director, Professor S.V. Egerev, and by the Deputy Director, Professor V.I. Mazepov. The Technological Department of the institute is headed by Chief Engineer N.S. Isaev. The Northern Branch is headed by Candidate of Physics and Mathematics S.P. Aksenov, a former student of Professor V.Yu. Zavadskii.

The basic research and engineering developments are carried out at 22 laboratories, which are headed by the leading scientists of the institute: Yu.Ya. Borisov, O.P. Galkin, E.P. Gulin, A.V. Evtushenko, I.B. Esipov, V.P. Kodanev, V.M. Lekomtsev, M.A. Margulis, V.Yu. Machnev, M.A. Mironov, V.A. Pirogov, R.Yu. Popov, P.A. Pyatakov, V.D. Svet, A.N. Serebryanyi, V.I. Sizov, V.N. Telepnev, Yu.I. Tuzhilkin, V.V. Tyutekin, and D.P. Frolov. The laboratories of the Northern Branch are headed by V.A. Zhuravlev and A.N. Neronov.

The institute takes part in basic and applied research under international, federal, and industrial programs. One of the main aspects of its activity is the work on its own scientific and technological projects. Since 1999, the institute, as a State research center, has carried out basic research in the fields of science specified by the Federal Scientific-Engineering Program "Research and Development in the Priority Directions of Civil Science and Engineering."

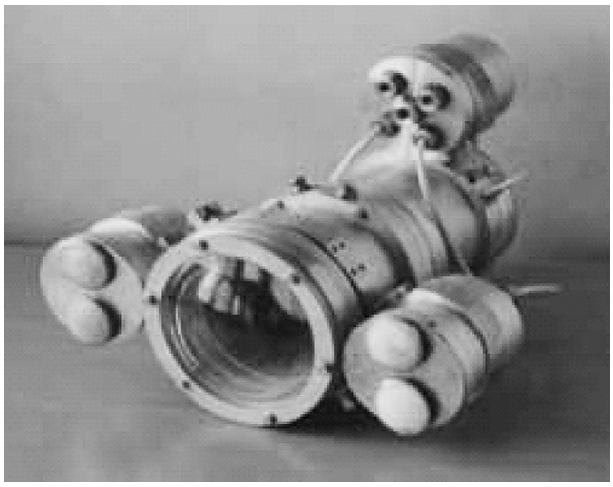
Underwater Acoustics. Along with continuing comprehensive studies of underwater sound propagation for naval purposes, the Acoustics Institute carries out systematic research in underwater acoustics. This research is aimed at the refinement of acoustic models of the ocean and seas and at the development of methods for predicting the energy and statistical characteristics of sound fields in application to specific oceanological conditions. Data obtained from previous expeditions in seas and oceans are collected to form acoustic oceanographic data bases. Methods of the ocean zoning according to these data are developed. An important applied aspect of the work on ocean and sea acoustics is the development of the marine test stations at the Black Sea and of the new acoustic methods for underwater observation of the variability of ocean and sea regions (O.P. Galkin, R.F. Shvachko, K.D. Sabinin, and A.N. Serebryanyi). Methods and means for the acoustical monitoring of the ocean, seas, and internal water bodies of Russia are developed to control the environment and to predict anomalous natural phenomena and antropogenic catastrophic events. Studies are carried out and methods are developed to detect and determine the epicenters of underwater earthquakes and to predict tsunamis on the basis of analyzing the hydroacoustic fields in the ocean (O.P. Galkin). Acoustic methods of gas hydrate–deposit exploration are developed, and engineering approaches to constructing the acoustic means for such explorations on the North Sea shelf are justified (I.B. Esipov).

Sources and Receivers of Sound. In the last few years, new principles of designing parametric radiators and receivers of sound on the basis of nonlinear acoustic phenomena have been widely investigated. The experimental studies performed with parametric arrays for ranges up to several hundreds of kilometers in the ocean showed the good prospects for the practical application of such sources (I.B. Esipov). A fundamentally new direction in the development of sound sources is the recent study of efficient low-frequency transmitters for ocean investigations. The tests of the pilot models of such sources showed that they are rather promising and deserve further development (V.A. Pirogov, I.A. Urusovskii, and M.A. Mironov). The search for new electromechanically active materials (composite, elastomer, liquid-crystal, etc.) and their study are carried out to provide the basis for a further development of the methods of electroacoustic energy transformation. New magnetostrictive alloys of rare-earth metals, electrostrictive ceramics, and composite and film-based elastic piezoelectric materials are investigated. The results of these studies are aimed at the development of a new generation of efficient acoustic sources and receivers with preset parameters, which are necessary both for underwater acoustics and for the intensification of different technological processes, geological surveys, medicine, and environmental monitoring (V.A. Pirogov, I.P. Golyamina, V.D. Svet, and V.I. Sizov).

Acoustic Signal Processing. Computer models of signals and noise are developed along with the methods and algorithms for solving a wide class of problems, including the detection, identification, and localization of various objects. A software that implements the modern adaptive algorithms of data processing on the basis of tomographic investigations is elaborated (E.P. Gulin).

Atmospheric Acoustics. On the basis of monitoring the sound and infrasound fields produced in the atmosphere by natural, industrial, or transportation sources, the physical models of these fields are refined. The physical principles of protection from infrasound radiation, which has the most adverse effect on humans, are justified (A.V. Evtushenko).

Methods and Means for Noise and Vibration Control. Physical foundations and principles of suppressing noise and vibration caused by mechanisms and machines in transportation and industry are developed. Computer programs are elaborated for determining the characteristics of noise produced by engineering structures. Methods for the optimum combination of different measures for reducing the noise and vibra-



Ultrasonic imaging system: an instrument for observation in a turbid medium. A new development of the Acoustics Institute.

tion of equipment are justified. The studies are aimed at increasing the service life of machines and mechanisms and improving the conditions of work for personnel (V.Yu. Machnev, M.A. Mironov, A.A. Dogadov, and A.I. Orlov). Principles of designing acoustic baffles for reducing traffic noise are developed (O.V. Kudryavtsev and I.A. Urusovskii).

Photoacoustics. Basic laws governing the formation of acoustic signals due to the effect of laser radiation on an inhomogeneous medium are determined. Photoacoustic methods for nondestructive testing and determination of the physical properties and parameters of materials are developed. The subject of main interest is the photoacoustic testing of special solutions and suspensions for medical purposes, for example, lipopolysaccharide solutions and nanoparticle suspensions (S.V. Egerev and A.V. Fokin). Photoacoustic methods are successfully used for analyzing the physical properties of carbon nanotubes and nanostructured amorphous materials (I.A. Chaban and O.B. Ovchinnikov).

Nonlinear Acoustics. Nonlinear acoustic phenomena that occur in inhomogeneous structures, granular media, and rock are studied. The results of the studies are aimed at developing physical models of acoustic impact on the efficiency of oil and gas production (I.B. Esipov).

Sonochemistry. Physical–chemical effects caused by acoustic cavitation in low-frequency acoustic fields, namely, emulsification, sonoluminescence, and dispersion of substances, are investigated. The principles of controlling the parameters of intense low-frequency cavitation fields are justified for applying them in industry and in solving some environmental problems (M.A. Margulis).

Bioacoustics. Neural mechanisms of auditory analysis and methods of data processing are investigated on the basis of bioacoustic analogies. Neural mechanisms that allow the auditory system to analyze complex signals in a wide dynamical range and in the presence of noise are revealed. Mechanisms of sounding signal radiation in dolphins are modeled. Principles of modeling the mechanisms of spatial hearing are studied. The studies are aimed at the development of experimental systems for testing machines and mechanisms, high-quality devices for hearing aids, devices for improving speech intelligibility, and new means of acoustic signal processing on the basis of cybernetic methods and bioacoustic analogies (N.G. Bibikov, N.A. Dubrovsky, and V.N. Telepnev).

Medical Acoustics. Medical–biological foundations of the application of ultrasound in medicine are developed, including methods of acting with focused ultrasound on the human organism and on human biological tissues. For this purpose, focusing (and also multifocus) radiators with a controlled structure of their ultrasonic fields are designed. The results of the studies should make it possible to create new medical instruments for surgery and therapy and to extend the application of ultrasound to the diagnosis and treatment of various disorders (L.R. Gavrilov and P.A. Pyatakov).

Applied Projects. Among the applied projects carried out under the Federal Program "Research and Development in the Priority Directions of Civil Science and Engineering," here we note two projects: first, the development of a series of pneumoacoustic atomizers for fine-disperse atomization of liquids in fire-control systems in closed rooms, for efficient burning of liquid fuel in steam boilers and heating plants, and for fine-disperse atomization of liquid solutions used in agriculture (Yu.Ya. Borisov); second, the development of ultrasonic imaging systems of different types, including those lowered from the board of a ship and self-contained ones for scuba divers, to provide for a rapid survey of underwater structures and to make it possible to carry out technological operations in turbid water or on a silt bottom (V.I. Sizov).

Federal Special-Purpose Program "National Technological Base." Since 2002, the Acoustics Institute has taken part in the development of critical acoustic technologies under the Federal Special-Purpose Program "National Technological Base." As a result, an acoustic chamber for underwater observations at depths up to 300 m was developed; a high-speed system of long-range underwater communication was designed; a parametric acoustic system for underwater monitoring of a shallow sea was elaborated; and structures of high-efficiency noise-absorbing materials were invented (V.P. Kodanev, V.Yu. Machnev, and V.I. Sizov).

CONCLUSIONS

During the fifty-year history of the Acoustics Institute, its scientists have taken an active part in the formation of the promising problems and directions of

research in acoustics that were developed in our country. Despite the fact that, more than 40 years ago, the institute was formally transferred from the system of research institutes of the Academy of Sciences to another system, its leading specialists continued to participate in the work of various commissions of the Academy of Sciences. The Scientific Council on Acoustics of the Academy of Sciences has traditionally been chaired by scientists from the Acoustics Institute (A.V. Rimskii-Korsakov, L.M. Lyamshev, and S.V. Egerev). Following the principles of the continuous popularization and development of scientific achievements in acoustics all over the country, the scientists of the institute initiated the establishment of the Russian Acoustical Society. Many specialists from the Acoustics Institute have become active members. The Russian Acoustical Society was established in June 1991, at the last (11th) All-Union Acoustical Conference, and now has its own history. Today, the Society (its president is Academician of the Russian Academy of Natural Sciences N.A. Dubrovsky, and its executive director is E.V. Yudina) has more than 600 members and extends its activities to 49 regions of Russia. These activities are coordinated by five territorial branches of the society: the Far East Branch (headed by Academician of the Russian Academy of Sciences V.A. Akulichev), the South Branch (Academician of the Russian Academy of Natural Sciences V.I. Timoshenko), the Siberian Branch (Academician of the Russian Academy of Natural Sciences V.K. Kedrinskii), the Volga Branch (Professor S.N. Gurbatov), and the Moscow Branch (Corresponding Member of the Russian Academy of Sciences O.V. Rudenko). Members of the society include researchers, professional workers, directors of enterprises and institutes, deans and chairs of departments of universities, university professors, and postgraduate and graduate students from more than 70 organizations and more than 60 cities of Russia and Ukraine.

An important part of the activity of the Russian Acoustical Society and the Acoustic Institute is the organization of the annual sessions where the current status of research in acoustics in Russia is discussed. Today, the sessions of the Russian Acoustical Society approach the previous All-Union Acoustical Conferences in the width of their scope and in the number of participants. As a member of the European Acoustics Association and a member of the International Institute of Noise-Control Engineering, the Russian Acoustical Society allows its members to present the results of their studies in the information structures of these international organizations. Currently, President of the Russian Acoustical Society N.A. Dubrovsky represents Russia in the International Commission for Acoustics.

For several years, the Russian Acoustical Society has held a competition among Russian graduate and postgraduate students for scholarships from the American Acoustical Society in support of the most interesting projects carried out by young researchers beginning their career in science. By now, 60 graduate and postgraduate students from different universities and research institutes have already received such scholarships. The geography of the origins of these young scientists actually reproduces the geography of our country: from St. Petersburg to Vladivostok and from Arkhangel'sk to Taganrog.

The Acoustics Institute, together with the Russian Foundation for Basic Research, supports the publishing activities of the Russian Acoustical Society. In addition to the proceedings of the sessions of the Russian Acoustical Society, the Annuals of the Russian Acoustical Society have appeared in the last few years. These annuals basically consist of series of papers presenting the results of studies devoted to a single area of research, namely, "Acoustics of Inhomogeneous Media." Most of the papers of this series were presented and discussed at the regular seminar held by Professor S.A. Rybak at the Acoustics Institute. Owing to the advanced level of problems discussed at this seminar, it has become widely known as a special kind of scientific club for discussing the problems of acoustics of inhomogeneous media. In recent years, a number of leading specialists in physics and acoustics of inhomogeneous media presented their results at this seminar. It should be noted that this area of research is rapidly progressing due to the efforts of many research groups in our country, and it is of interest to scientists specializing in solving both theoretical and experimental problems with different areas of application. In addition, the atmosphere of free discussion that is characteristic of the aforementioned seminar also attracts specialists working in adjacent areas of research. Useful information about current publications and the development of acoustic studies in Russia can be obtained on the web sites of the Acoustics Institute and the Russian Acoustical Society: www.akin.ru and www.akin.ru/rao.

In the framework of a single paper, it is difficult to describe the results of the great creative work of hundreds of people working at the Acoustics Institute—professors, doctors and candidates of science, engineers, technicians, and laboratory assistants—all those who, within 50 years, have laid the scientific foundations for modern acoustics and hydroacoustics. Many of these specialists, who possess priceless experience, are still working at the institute and creating new acoustic technologies. More fully than in this paper, the many-sided activities of the institute are represented by the publications of our scientists. A comprehensive description of the life and history of the institute can be found in the jubilee book devoted to the 50th anniversary of the Acoustics Institute, which currently is in active preparation for publication. The list of references presented below contains the monographs written by authors from the Acoustics Institute and published

within the 50 years of its existence and also two papers devoted to the history of the institute.

REFERENCES

1. V. A. Akulichev, *Cavitation in Cryogenic and Boiling Liquids* (Nauka, Moscow, 1978) [in Russian].
2. A. A. Anan'eva, *Ceramic Receivers of Sound* (Akad. Nauk SSSR, Moscow, 1963) [in Russian].
3. I. B. Andreeva and V. G. Samovol'kin, *Scattering of Acoustic Waves by Marine Organisms* (Agropromizdat, Moscow, 1986) [in Russian].
4. I. B. Andreeva, *Physical Fundamentals of Sound Propagation in the Ocean* (Gidrometeoizdat, Leningrad, 1975) [in Russian].
5. V. M. Bel'kovich and N. A. Dubrovskiĭ, *Sensory Basis of Cetacean Orientation* (Nauka, Leningrad, 1976; U.S. Joint Publication Research Service, 1977).
6. N. G. Bibikov, *Description of Sound Features by Neurons of the Auditory System of Terrestrial Vertebrates* (Nauka, Moscow, 1987) [in Russian].
7. L. M. Brekhovskikh, *Waves in Layered Media*, 2nd ed. (Nauka, Moscow, 1973; Academic, New York, 1980).
8. L. M. Brekhovskikh and Yu. P. Lysanov, *Fundamentals of Ocean Acoustics* (Gidrometeoizdat, Leningrad, 1982; Springer, New York, 1991).
9. I. A. Viktorov, *Rayleigh and Lamb Waves: Physical Theory and Applications* (Nauka, Moscow, 1966; Plenum, New York, 1967).
10. I. A. Viktorov, *Surface Sound Waves in Solids* (Nauka, Moscow, 1981) [in Russian].
11. V. I. Volovov, *Sound Reflection from an Oceanic Bottom* (Nauka, Moscow, 1993) [in Russian].
12. L. R. Gavrilov and E. M. Tsirul'nikov, *Focused Ultrasound in Physiology and Medicine* (Nauka, Leningrad, 1980) [in Russian].
13. G. V. Glekin, *Nikolaĭ Nikolaevich Andreev* (Nauka, Moscow, 1980) [in Russian].
14. N. A. Dubrovskiĭ and V. I. Mazepov, in *Some of the History of National Hydroacoustics* (St. Petersburg, 1998), pp. 86–113 [in Russian].
15. V. Yu. Zavadskiĭ, *Calculation of Wave Fields in Open Regions and Waveguides* (Nauka, Moscow, 1972) [in Russian].
16. V. Yu. Zavadskiĭ, *The Finite-Difference Method in Wave Problems of Acoustics* (Nauka, Moscow, 1982) [in Russian].
17. A. V. Il'in, *Geomorphology of the Atlantic Ocean Bottom* (Nauka, Moscow, 1976) [in Russian].
18. A. V. Il'in, V. V. Orlenok, and I. I. Shurko, *Petrophysics of the Ocean Bottom*, Ed. by G. B. Udintsev (Mezhduved. Geofiz. Kom. Ross. Akad. Nauk, Moscow, 1992) [in Russian].
19. M. A. Isakovich, *General Acoustics* (Nauka, Moscow, 1973) [in Russian].
20. I. F. Kadykov, *Acoustics of Underwater Earthquakes* (Nauka, Moscow, 1986) [in Russian].
21. V. F. Kazantsev, L. D. Rozenberg, and O. L. Makarov, *Ultrasonic Cutting* (Akad. Nauk SSSR, Moscow, 1962) [in Russian].
22. A. P. Kapustin and O. A. Kapustina, *Acoustics of Liquid Crystals* (Nauka, Moscow, 1986) [in Russian].
23. K. V. Konyaev and K. D. Sabinin, *Waves within the Ocean* (Gidrometeoizdat, St. Petersburg, 1992) [in Russian].
24. L. M. Lyamshev, *Sound Reflection from Thin Plates and Shells in a Liquid* (Akad. Nauk SSSR, Moscow, 1955) [in Russian].
25. L. M. Lyamshev, *Laser Thermooptical Excitation of Sound* (Nauka, Moscow, 1989) [in Russian].
26. L. M. Lyamshev, *Radiative Acoustics* (Nauka, Moscow, 1996) [in Russian].
27. K. A. Naugol'nykh and L. A. Ostrovskiĭ, *Nonlinear Wave Processes in Acoustics* (Nauka, Moscow, 1990) [in Russian].
28. K. A. Naugol'nykh and N. A. Roĭ, *Electrical Discharges in Water* (Nauka, Moscow, 1971) [in Russian].
29. V. V. Ol'shevskiĭ, *Statistical Methods in Underwater Detection and Ranging* (Sudostroenie, Leningrad, 1983) [in Russian].
30. A. V. Rimskiĭ-Korsakov, *Electroacoustics* (Svyaz', Moscow, 1973) [in Russian].
31. A. V. Rimskiĭ-Korsakov, D. V. Bazhenov, and L. A. Bazhenova, *Physical Principles of Sound Formation in Air-Blowing Machines* (Nauka, Moscow, 1988) [in Russian].
32. Yu. M. Sukharevskiĭ, in *Some of the History of National Hydroacoustics* (St. Petersburg, 1998), pp. 114–123 [in Russian].
33. L. A. Chernov, *Wave Propagation in a Random Medium*, 2nd ed. (Nauka, Moscow, 1975; McGraw-Hill, New York, 1960).
34. *Acoustics of the Ocean*, Ed. by L. M. Brekhovskikh (Nauka, Moscow, 1974) [in Russian].
35. *Acoustics of Turbulent Flows*, Ed. by A. V. Rimskiĭ-Korsakov (Nauka, Moscow, 1983) [in Russian].
36. *Acoustic Aerohydrodynamic Studies*, Ed. by A. V. Rimskiĭ-Korsakov (Nauka, Moscow, 1975) [in Russian].
37. *Aeroacoustics*, Ed. by A. V. Rimskiĭ-Korsakov (Nauka, Moscow, 1980) [in Russian].
38. *Vibrations and Noise (Physical Research)*, Ed. by A. V. Rimskiĭ-Korsakov (Moscow, 1969) [in Russian].
39. *Vibrations, Radiation, and Damping of Elastic Structures*, Ed. by A. V. Rimskiĭ-Korsakov (Moscow, 1973) [in Russian].
40. *Ultrasound. A Small Encyclopedia*, Ed. by I. P. Golyamina (Sovetskaya Éntsiklopedia, Moscow, 1979) [in Russian].
41. *Physics and Technology of High-Intensity Ultrasound*, Vol. 1: *Sources of High-Intensity Ultrasound*, Vol. 2: *High-Intensity Ultrasonic Fields*, Vol. 3: *Physical Fundamentals of Ultrasonic Technology*, Ed. by L. D. Rozenberg (Nauka, Moscow, 1967, 1968, 1970) [in Russian].

Translated by E. Golyamina

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 33–37. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 47–51.

Acoustic Field Generated by a Beam of Protons Stopping in a Water Medium

V. I. Albul*, V. B. Bychkov*, S. S. Vasil'ev*, K. E. Gusev, V. S. Demidov**, E. V. Demidova**, N. K. Krasnov*, A. F. Kurchanov*, V. E. Luk'yashin**, and A. Yu. Sokolov****

** All-Russia Research Institute for Physical-Technical and Radiotechnical Measurements, Mendeleevo, Moscow oblast, 141570 Russia*

*** State Research Center Institute of Theoretical and Experimental Physics, ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117259 Russia*

e-mail: Demidov@itep.ru

Received October 27, 2003

Abstract—In an experiment with a beam of protons accelerated up to an energy of 200 MeV, the space–time structure of the hydroacoustic field generated by protons stopping in the water medium was observed. The contributions of three components were separated: a cylindrical wave diverging from the middle part of the acoustic antenna and the signals from the ends of the antenna, namely, from the region of the maximal energy density release by protons (the Bragg peak) and from the other end corresponding to the beam entrance into water. © *2005 Pleiades Publishing, Inc.*

The properties of acoustic radiation generated in the passage of ionizing particles through matter have been studied theoretically and experimentally for several decades [1–11]. Currently, its main characteristics are known: the proportionality of the response of acoustic transducers to the intensity of the proton beam, the interdependence between the diameter of the beam and the duration of the response signal, the temperature dependence of the intensity of an acoustic signal arising in liquids, and so on. All experimental data were obtained using proton and electron accelerators. In the experiments, the radiation from the nearest point of the antenna was measured without analyzing the acoustic field produced by other parts of the antenna. Further experiments in this area, including the measurement of the space–time structure of the field, are important not only for determining the radiation mechanisms but also from the point of view of its application, for instance, to the problem of detection of ultra-high-energy cosmic neutrinos in natural water basins [12–14]. Hydroacoustic methods of detecting the cascades caused by cosmic particles have a number of advantages compared to the conventional scintillation method. These methods include the low absorption of sound in water and its much lower velocity compared to the same parameters of light. This leads to a reduction in the experiment cost and makes it possible to use a great number of hydrophones in realizing the acoustic method of cosmic particle detection.

The current status of the problem of diagnosing the ionizing radiation by hydroacoustic methods is described in [15, 16]. A successful practical application of sound waves in solving the problem of measuring the spectra of cosmic neutrinos requires not only the detection of the desired signal against the background of sea noise but also the determination of no less than four parameters characterizing the source of radiation by using the detected signal: the coordinates and size of the radiating area, the direction of propagation, and the energy of a cosmic particle. The study of sea noise is a separate problem; the objective of the present paper is to study, under laboratory conditions, the acoustic field generated due to the deceleration of an intense beam of accelerated protons in water down to a stop, which is the best model of a nuclear-electromagnetic cascade.

A beam of accelerated ionizing particles, axially symmetric along the direction of propagation, upon entering an initially homogeneous boundless equilibrium liquid practically instantly creates a heated zone in a limited space region. The process of heating can be conveniently described by a distribution of instantly acting heat sources in a cylindrical coordinate system connected with the beam, with the Z axis directed along the propagation of the beam of particles. As a result, perturbations of the pressure, density, and temperature fields appear in the initially stationary liquid. The intensity of the perturbations is proportional to the specific energy density loss in the beam of particles, which has a noticeable maximum at the last two centimeters of the particle path. Figure 1 shows a double differential distribution of the mean energy loss of protons in water, which was obtained by modeling with the GEANT-3.21 program [17]. The initial proton energy was 200 MeV. The upper part of Fig. 1 displays the contours of the

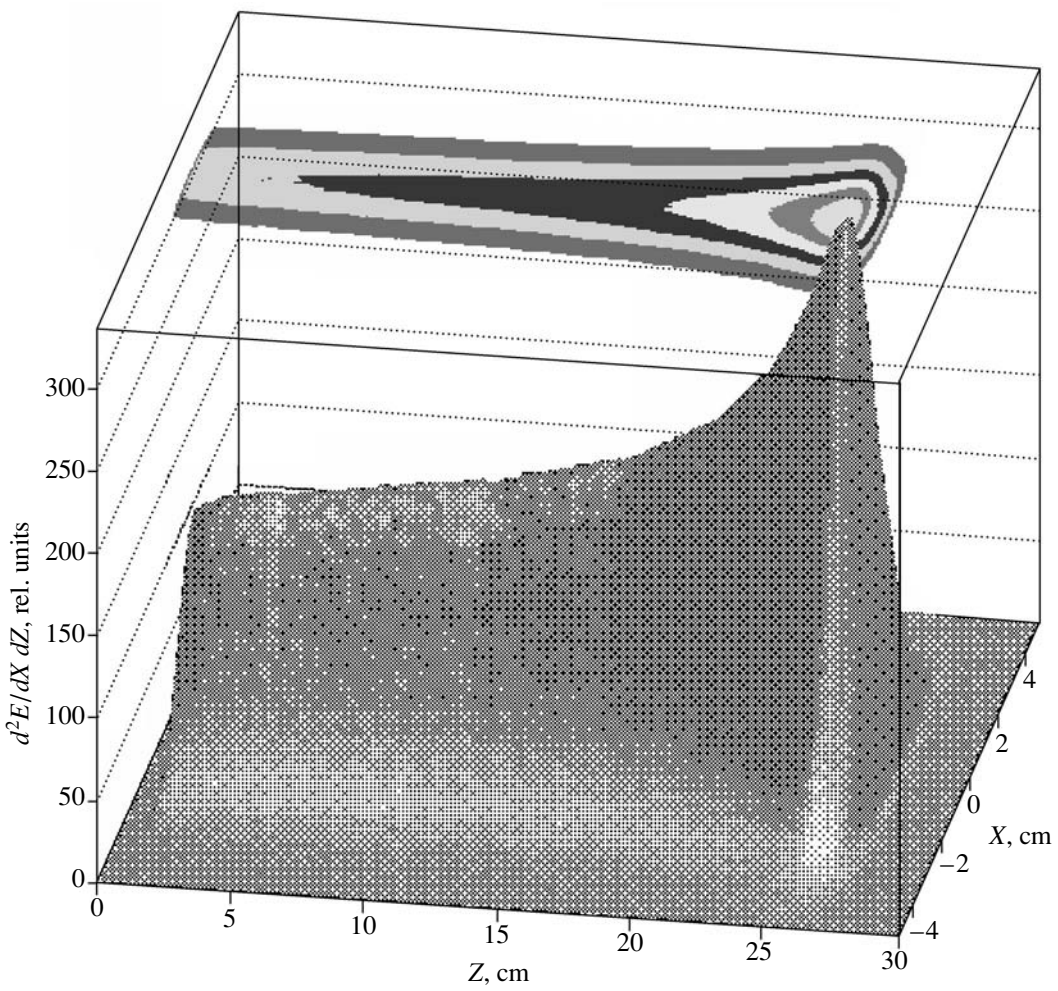


Fig. 1. Double differential distribution of the mean energy loss of protons in water, which was obtained by modeling. The protons propagate along the *Z* axis, and the *X* axis lies in the horizontal plane passing through the beam center. The beam energy is 200 MeV. In the upper part of the plot, the contours of the two-dimensional distribution of energy loss are shown for the levels of 0.85, 0.7, 0.56, 0.42, 0.28, and 0,14 of the maximum.

two-dimensional loss distributions for the levels of 0.85, 0.7, 0.56, 0.42, 0.28, and 0.14 of the maximum. A similar spatial form has an antenna that appears at the instant of the beam passage and creates the acoustic field in the medium.

The experiment to measure the parameters of the hydroacoustic field of a proton beam of charged particles was carried out using an external proton beam at the accelerator of the Institute of Theoretical and Experimental Physics that is similar to the accelerator employed in [4]. The beam had the following parameters: the energy was 200 MeV, the beam pulse duration was 70 ns, the spatial form of the beam in the transverse direction was close to a Gaussian distribution, and the beam diameter was 2 cm (at a half-maximum level). The beam intensity was maintained constant at about 4×10^{10} protons per pulse and was monitored by an inductive sensor. As a target, in which the proton beam was decelerated, we used salt water (the salt concentration was about 3%) filling a tank (Fig. 2) in the form of a parallelepiped $50.8 \times 52.3 \times 94.5$ cm in size. The tank was made of acrylic plastic and was reinforced and sealed at the side joints. It was filled 90% full. The water temperature was 13.5°C. The size and equipment of the target made it possible to prevent the reflection from the boundaries during the measurements and obtain a detailed structure of the hydroacoustic field of a proton beam. The injection of the beam at the center of the measurement volume was carried out through a tube of diameter 59 mm and length 46 cm with a wall thickness of 1.5 mm, which was inserted in a lateral face of the tank and was closed with a teflon cap 2 mm thick. The mean proton path length in water was 25.2 cm. Other sizes characterizing the relative positions of equipment in the experiment are given in Fig. 2.

The measurements were performed using two GI-14 hydrophones with a transducer diameter of 4 mm and a height 6 mm. One of the hydrophones, H2, was fixed

and the position of the other, H1, was changed during the experiment. The GI-14 hydrophone has a linear amplitude–frequency characteristic in a frequency range greater than 100 kHz, which makes it possible to record an undistorted acoustic pulse generated in water by the beam of protons. The hydrophone H1 was moved using a specially designed electromechanical remote-control scanner, which enabled us to position the hydrophone discretely, at a step of 8.9 mm, within a linear aperture 0.4 m long.

The signals from the hydrophones and the inductive sensor of beam intensity were entered into a standard IBM PC using a specially developed board. The analog part of the input board allows one to simultaneously digitize up to four analog signals with a maximum sampling frequency of 400 kHz for every channel. The read-out of information was carried out by the signal fed to the sync input of the board simultaneously with the arrival of the beam at the target. A special system software was developed for this board, which provided a simultaneous input and displaying and recording the signals of all channels with synchronization by any signal. A set of application software for signal processing was also developed, which made it possible to compute the amplitude–phase dependences of signals, the narrow-band spectrum, the third-octave spectrum, three-dimensional representations of the spatial and time dependences, and other signal characteristics.

The measurements of the acoustic field generated by the proton beam were carried out via scanning with the hydrophone H1 in the horizontal plane passing through the axis of the proton beam, along linear paths at distances of 3, 6, and 13.5 cm from the beam center (the 01, 02, and 03 runs of measurements, respectively). The coordinate Z_{H1} (along the beam) was varied from the point of entrance of the beam into water up to 36 cm. The time of realization for every position of the movable hydrophone was 250 μs (1000 readings of 0.25 μs each). The mathematical processing was performed using the aforementioned equipment and included frequency filtering, amplitude analysis and analysis of the waveform of the signal, construction of averaged data samplings for arbitrary coordinate directions, and three-dimensional imaging.

Figure 3 presents in a three-dimensional space–time structure of the hydroacoustic field produced by a proton beam introduced in the water target (the results of run 02 are illustrated in this example). The time sweep t is represented by the abscissa axis; the quantity Z_{H1} in scanner step units (p1–p40), by the ordinate axis; and the voltage at the output of the hydrophone H1 in relative units, by the Z axis. For better comprehension of the field structure, the signal is inverted in the figure. The dynamics of variation of hydrophone response can be seen. The contributions from three coherent sources can be separated. The acoustic signal from the nearest point of the radiating acoustic antenna is observed as an extended valley in the A–B direction (the first source).

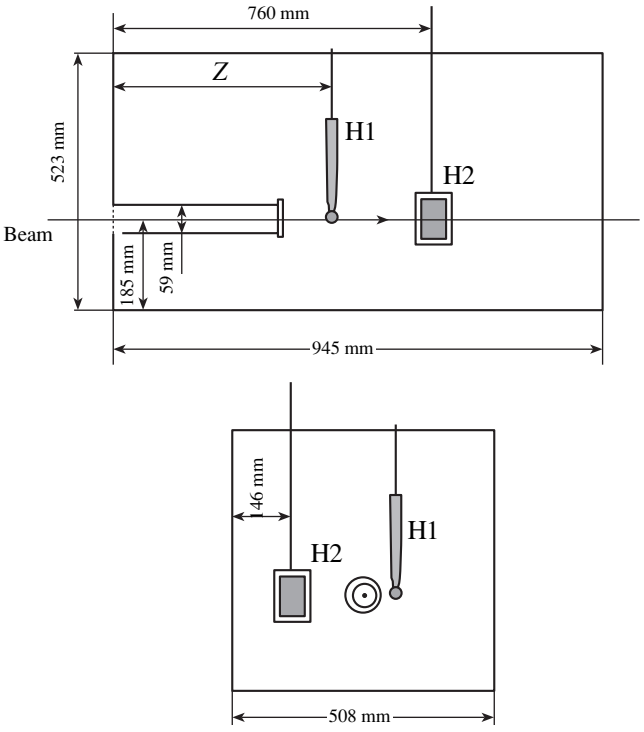


Fig. 2. Schematic diagram of the hydrohone positions in the water target.

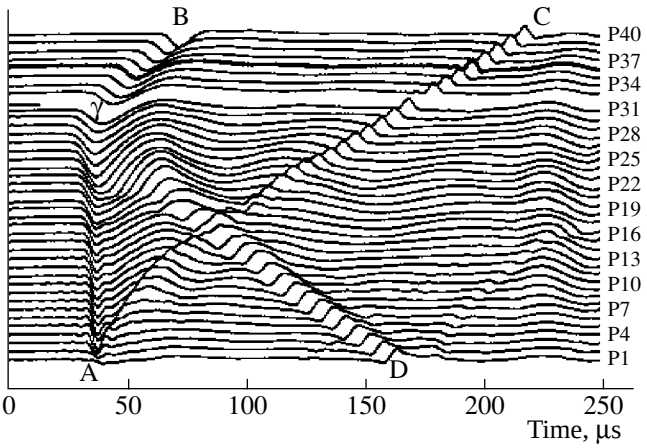


Fig. 3. Space–time structure of the acoustic field generated in water by a beam of protons with an energy of 200 MeV.

The longitudinal size of the antenna corresponds to the proton path length in water. In Fig. 3, the letter γ denotes the region of termination of the beam and, correspondingly, the end of the acoustic antenna. In the initial part (before the region γ), the valley is parallel to the ordinate axis, because the hydrophone moves parallel to the antenna and the time of signal propagation to the hydrophone remains constant. In this part, a cylindrical wave is detected, which, according to theory [6], diverges from the antenna. Further, the hydrophone, continuing its movement along a straight line, recedes

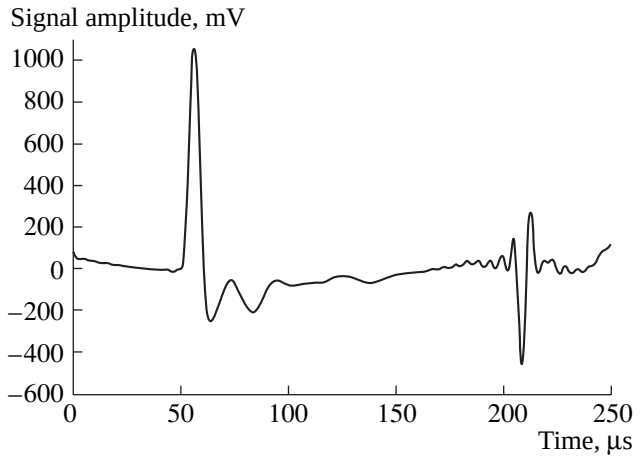


Fig. 4. Waveform of the acoustic signal.

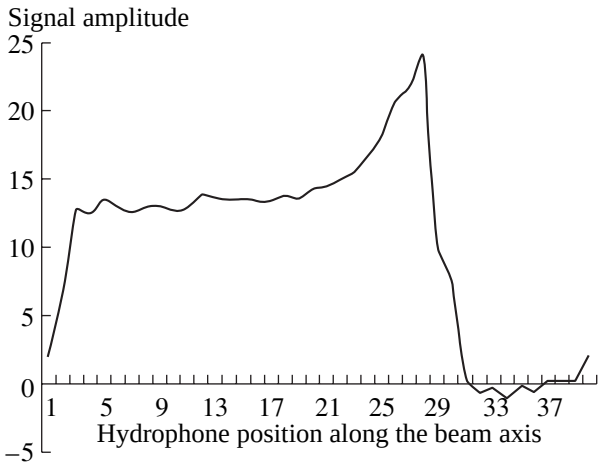


Fig. 5. Acoustic field in the near-field zone.

from the antenna, the time of signal arrival increases, and the valley in the section γ–B changes its direction. Note that the main energy of the acoustic signal from this source manifests itself as a compression wave.

The other source, the signal from which is presented in Fig. 3 by the ridge DB, is the area of the so-called Bragg peak (at the end of the proton path), where, according to Fig. 1, the energy density release rises sharply. When the hydrophone is located at the beginning of the antenna, near the point of the entrance of the beam into the water (the curve corresponding to this case is designated as P1 in Fig. 3), the area of the Bragg peak is spaced from the hydrophone by 25.2 cm and the signal from this area arrives within 168 μs after the passage of the beam (near the point D). As the hydrophone moves along the beam, the Bragg peak approaches the hydrophone, and the corresponding branch of the ridge DB moves from right to left. As the γ area is approached, the interference of the signals AB and DB and their junction in the section γB are observed. It is significant that the signal from this source appears in the section Dγ mainly as a rarefaction wave; then (γB), the signal changes its polarity, and a wave of increased pressure propagates in this direction.

The third source, the signal from which is shown in Fig. 3 as a ridge AC, corresponds to the cap through which the proton beam enters the tank. As the hydrophone moves along the beam, the source recedes from the hydrophone and the source trajectory is directed from left to upward right. The bending of the trajectory at the beginning takes place due to the interference of this signal with the signal from the first source. The polarity of this signal coincides with the polarity of the signal from the Bragg peak in the section Dγ.

Figure 4 displays the waveform of a typical acoustic pulse P39 from run 01 in a direct (not inverted) form. The curve is recorded at a point, the *z* coordinate of which is 8.3 cm greater than the coordinate of the Bragg peak area. The first signal detected by the hydrophone comes from the area of the beam stopping. It is characterized by a sharp increase in pressure, and its duration is 10.5 μs. The following several periods of relaxation have considerably less amplitudes and attenuate within approximately 70 μs. The second signal comes from the third source. In its shape, it looks like a signal from a laser thermoacoustic sound source in the far-field zone (see Ch. 10 of [6]) and is characterized by alternating signals of increased and reduced pressure with different amplitudes. On the whole, the energy of this signal is the rarefaction and its duration is 4.5 μs.

Figure 5 shows the distribution of acoustic pressure in the near-field zone of the antenna: the dependence of the signal amplitude on the hydrophone position along the beam axis on the path nearest to the beam (measurement 01). The shape of this distribution coincides with the well-known shape of the ionization loss curve. The drop of the curve at the beginning is explained by the fact that, at the first two points, the hydrophone was located before the point of the entrance of the beam into water. The sharp rise at the end of the curve corresponds to the so-called Bragg peak, i.e., the increase in the ionization loss at the end of the particle path in the matter, which is observed at a distance of 23.8 cm from the point of entrance of the beam into water. The calculated position of the Bragg peak for a proton beam of such an energy is 25.2 cm [17]. The difference is possibly connected with the interference of the signals AB and DB in the area γ.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 02-02-17148a, and by the Ministry of Science, grant NSh-1867.2003.2.

REFERENCES

1. G. A. Askaryan, At. Energ. **3** (8), 152 (1957).

2. G. A. Askariyan, B. A. Dolgoshein, A. N. Kalinovsky, and N. V. Mokhov, Nucl. Instrum. Methods **164**, 267 (1979); G. A. Askariyan, At. Energy **3**, 152 (1957).
3. R. M. White, J. Appl. Phys. **34**, 2123 (1963).
4. L. Sulak, T. Armstrong, H. Baranger, *et al.*, Nucl. Instrum. Methods **161**, 203 (1979).
5. V. D. Volovik, in *Abstracts of VI International Symposium on Nonlinear Acoustics* (Moscow, 1975), p. 191.
6. V. I. Albul, V. B. Bychkov, K. E. Gusev, *et al.*, Prib. Tekh. Éksp., No. 3, 50 (2001) [Instrum. Exp. Tech. **44**, 327 (2001)].
7. V. I. Albul, V. B. Bychkov, K. E. Gusev, *et al.*, Izmer. Tekh., No. 8, 58 (2003).
8. V. D. Volovik and V. I. Kobizskiĭ, Pis'ma Zh. Tekh. Fiz. **2** (2), 66 (1976) [Sov. Tech. Phys. Lett. **2**, 25 (1976)].
9. S. D. Hunter and W. V. Jones, J. Acoust. Soc. Am. **69**, 1557 (1981).
10. L. M. Lyamshev, *Radiative Acoustics* (Nauka, Moscow, 1996) [in Russian].
11. A. I. Pushkarev, M. A. Pushkarev, and G. E. Remnev, Akust. Zh. **48**, 260 (2002) [Acoust. Phys. **48**, 220 (2002)].
12. L. G. Dedenko, I. M. Zheleznykh, S. Kh. Karaevskiĭ, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **61** (3), 593 (1997).
13. A. Butkevich and I. M. Zhelezhykh, in *Proceedings of 2nd NESTOR International Workshop* (Pylos, Greece, 1992), p. 345.
14. L. G. Dedenko, A. V. Furduev, Ya. S. Karlik, *et al.*, in *Proceedings of 25th International Cosmic Rays Conference, Durban, 1997* (World Sci., Singapore, 1997), Vol. 7, p. 89.
15. V. B. Bychkov, S. S. Vasilijev, and V. S. Demidov, in *Proceedings of XVI International Symposium on Nonlinear Acoustics ISNA-16* (Moscow, 2002), Vol. 2, p. 909.
16. V. I. Albul, V. B. Bychkov, K. E. Gusev, *et al.*, in *Proceedings of XX Symposium on Hydroacoustics* (Gdansk, Poland, 2003), p. 293.
17. *GEANT Detector Description and Simulation Tool (Long Writeup W5013)* (CERN, Geneva, 1995).

Translated by A. Svechnikov

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 38–43. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 52–58.

Algorithm for Calculating the Cross-Section Areas of the Vocal Tract

P. Badin, I. S. Makarov*, and V. N. Sorokin***

**Institute for Information Transmission Problems, Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, Moscow, 101447 Russia*

e-mail: speechprod_mak@mail.ru; vns@iitp.ru

***Institut de la Communication Parlée, Université Stendhal, INPG, Grenoble, UMR CNRS 5009, France*

e-mail: badin@icp.inpg.fr

Received February 15, 2003

Abstract—It is found that the articulation process is accompanied by active variations of the pharynx width. To describe the latter, a linear combination of two width eigenvectors with varying coefficients is proposed. A new algorithm is constructed for calculating the cross-section areas of the vocal tract. The algorithm takes into account not only the anatomic parameters and the shape of the tract in the saggital plane but also the parameters in the lateral and axial planes. *© 2005 Pleiades Publishing, Inc.*

In the problems of automatic recognition, synthesis, and compression of speech, it is expedient to determine the shape of the vocal tract from the parameters of the speech signal. Such a determination presents an inverse problem. To solve the inverse problem with respect to the shape of the vocal tract, it is necessary to have a "code book" that relates the measured acoustic parameters of the speech signal to the articulatory parameters. To create such a book, it is necessary to solve a number of particular problems and, first of all, to construct an algorithm for calculating the cross-section areas of the vocal tract from the current values of controlled articulatory parameters and also from the values of anatomic parameters of the vocal tract.

In the theory of speech production, the problem of calculating the cross-section areas from the current controlled articulatory parameters is usually solved using one of three approaches. The first approach is based on the so-called $\alpha\beta$ model [1–3]. According to this model, the cross-section area of the vocal tract $S(x)$ for each x coordinate, which is measured along the tract's midline beginning from the glottis, is calculated as $S = \alpha h^{\beta}$, where $h(x)$ is the distance measured from moving to immovable surfaces of the vocal tract in the mid-saggital plane and α and β are some coefficients to be determined.

In the second approach, the cross-section area function is represented as a linear combination or a quadratic form of some eigenvectors obtained by the factor analysis of experimental data [4]. The main advantage of this approach is that the vocal tract is parametrized with a very small number of parameters. In both the first and the second approaches, sounds that have identical (or almost identical) saggital profiles but different tongue shapes in the lateral and/or axial planes cannot be distinguished. Another disadvantage is referencing to a specific speaker and/or to a specific speech material.

In the third approach [5], it is assumed that, for an effective calculation of the cross-section areas, it is necessary to take into account not only the saggital but also the lateral and axial parameters of the vocal tract.

The algorithm for calculating the cross-section areas from the articulatory parameters is based on a mathematical model of articulation. We used the articulatory model described in [5]. This model is determined by the following articulatory parameters: the height of the glottis, the coordinates of the tongue root, the coordinates of the tongue tip, the rotation angle of the lower jaw, the horizontal displacement of the rotation point of the lower jaw, five coefficients multiplying the eigenfunctions of the elastic deformations of the tongue, and the height of the lower lip, which makes a total of 13 controlled parameters. The analysis performed by us showed that the model should be complemented with two more parameters, namely, the coefficients multiplying the eigenfunctions that describe the variations of the pharynx width (see below).

We denote the vector of articulatory parameters determining the shape of the vocal tract at an instant t as $u(t) = (u_1(t), u_2(t), \ldots, u_{15}(t))$. Knowing $u(t)$, one can calculate the moving surfaces of the tract: the tongue surface, the lip surface, and the lower part of the pharynx in the saggital cross section. From the moving surfaces of the tract and from the shapes of the hard palate and teeth, one can calculate the distance between the moving and immovable surfaces of the vocal tract $h(x)$. In [5], an algorithm was proposed for calculating the cross-section areas in different regions of the vocal tract from the function $h(x)$ and also from some additional articulatory parameters. However, a comparison of the

cross-section areas calculated with this algorithm and the areas measured by magnetic resonance imaging (MRI) showed that the algorithm describes the relation between the articulatory parameters and the areas in some regions of the tract with insufficient accuracy, and the errors prove to be much greater than those characteristic of the MRI. Therefore, it has become necessary to reconsider the algorithm used for calculating the areas on the basis of magnetic resonance measurements.

Currently, magnetic resonance imaging is the main experimental technique used for measuring the cross-section areas of the vocal tract. The speaker is placed inside a magnetic system producing a magnetic field, which rotates the hydrogen spins in the vocal tract tissue through 90° and 180°. To perform the scanning, several plates oriented across the vocal tract are positioned on the head of the speaker. Inside the plates, radio-frequency coils are placed. The coils excite the hydrogen spins in each layer by high-frequency radio pulses and also detect the reflected nuclear magnetic signal. The reflected signal is subjected to a two-dimensional Fourier transformation, which generates a magnetic resonance image of each of the layers [6].

In this paper, we use the results of the measurements performed in Grenoble (Centre Hospitalier Regional Universitaire de Grenoble, France) by one of the authors [7, 8]. The shape of the cross section of the vocal tract was studied for ten French vowels and six Swedish consonants. In each experiment, three sets of plates were used. The positions of these plates in the saggital plane are shown in Fig. 1. The first set of plates (from 1st to 18th) was placed across the larynx and the lower part of the vocal tract. These plates provided the cross-section areas of the larynx and the pharynx. The second set (from 19th to 37th) was positioned at an angle of 45° to the first set, across the middle part of the vocal tract. These plates provided the measurement of the cross-section areas of the tract in the region of the velum (plates 19 to 27). The third set (from 38th to 55th) was oriented at an angle of 90° to the first set, across the front part of the vocal tract. These plates provided the measurement of the cross-section areas of the tract in the region of the hard palate and alveoles. Each plate was divided into 65536 measurement units, i.e., pixels. The spatial resolution of each plate was 1 mm/pixel.

Owing to the difference in the anatomic features in different regions of the vocal tract, the dependences of the cross-section area on the tract's geometric dimensions are also different. We can separate six such regions (Fig. 2): (*1*) from the glottis to the esophagus inlet (the epilaryngeal region); (*2*) from the esophagus inlet to the projection of point *p* onto the back wall of the tract, where *p* is the center of the polar coordinate system in which the tongue surface is described; (*3*) from the projection of the center *p* onto the rear surface of the tract to the lower point on the velum; (*4*) the velum;

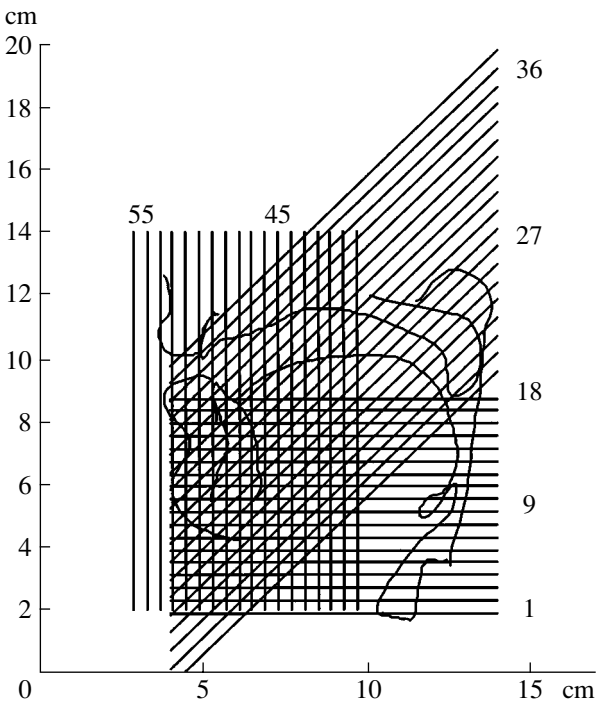


Fig. 1. Plate positions for magnetic resonance imaging in the saggital section. The total number of plates is 55. The thickness of each plate is 3.5 mm, and the distance between neighboring plates is 4 mm. The numbers in the plot indicate the order numbers of some of the plates (first plate, ninth plate, etc.).

(*5*) the hard palate (excluding the alveoles); and (*6*) the alveolar region. For the sixth region, the data were not obtained for all sounds. The lip region was not included in the analysis, because the shapes of the cross sections in this region were beyond the abilities of the MRI technique [7]. From the contours of the MRI cross sections, the following parameters were measured: the distance between the moving and immovable surfaces of the tract in the saggital cross section; the width of the tract (in the axial section) in its different parts; the transverse dimensions of the tongue, the uvula, and the lower jaw; and the thickness of the teeth. In addition to the shapes of the cross sections, the MRI also provided the saggital contours of the vocal tracts of both speakers for all sounds. An example of a saggital contour together with the boundaries dividing the vocal tract into six regions is shown in Fig. 2. The saggital contours were used to measure the shape of the hard palate, the coordinates of the tongue root and the center of the polar coordinate system *p*, the tongue surface, and the sublingual surface.

The pharynx has the form of a funnel-shaped channel, 7–10 cm long, with its wider end looking downwards. The cross section of the pharynx is determined by contractions of three muscles (the upper, middle, and lower pharynx constrictors). In [6, 8, 9], it was noted that, for different vowels, the width of the phar-

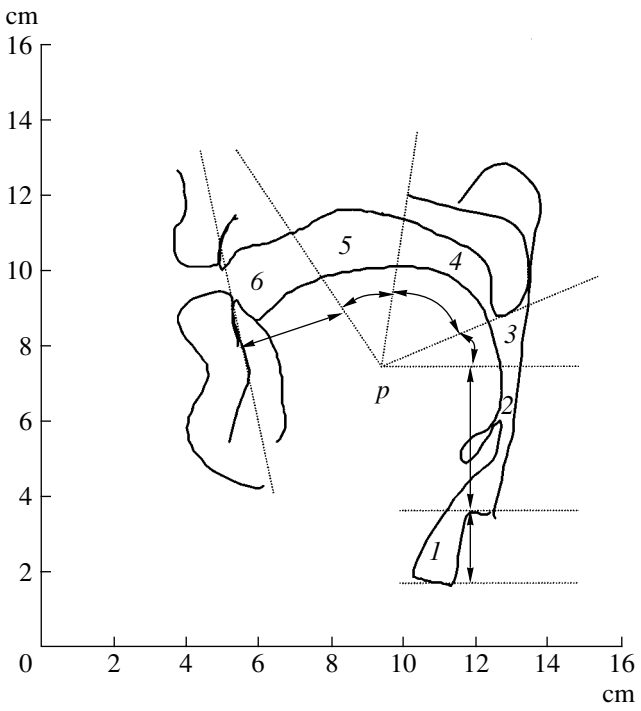


Fig. 2. Saggital MRI scan and six regions of the vocal tract.

ynx in the frontal plane as a function of coordinate along the tract exhibits different behavior. According to our data, the values of the pharynx width also were noticeably different for different vowels.

The variations of the pharynx width can be explained by the following factors:

(i) the muscle activity, i.e., the activity of the pharynx contraction muscles, which was observed in the electromiogram measurements for these muscles [10]; and

(ii) the rubber pipe effect: by compressing the pipe in one direction, we increase its transverse dimension in the other direction. This conclusion could be derived from the results reported in [11], where it was shown that the pharynx width approximately linearly depends on the saggital distance.

Studying the behavior of the pharynx width for all vowels in our experiments, we did not observe its dependence on the distance in the saggital plane, and the pharynx width could not be described by any analytical function of the saggital distance. This agrees well with the understanding of the activity of the pharynx contraction muscles in the speech production process.

The development of a mathematical model of the elastic deformations of the pharynx under the effect of contraction muscles encounters certain difficulties, and the efforts required to overcome them are not justified by the resulting accuracy of the solution. Instead of the direct mathematical approach, it is expedient to use a statistical analysis. The principal component method provides a sufficiently precise mathematical model of the pharynx width $P(x)$ on the basis of the experimental data analysis:

$$P(x) = \bar{P}(x) + \sum_i a_i f_i(x), \quad (1)$$

where x is the distance along the tract's midline, $f_i(x)$ is the ith eigenvector, and a_i is the factorial load. Two eigenvectors cover about 93% of the measurement variance. To these vectors, a certain physiological meaning can be ascribed. The first vector is associated with the activity of the upper pharynx constrictor, and the second, with the activity of the middle pharynx constrictor, while $\bar{P}(x)$ is the constant pharynx width characteristic of an individual speaker.

The function $\bar{P}(x)$ can be approximated by a hyperbola with a root-mean-square error of 3%:

$$y = 1/(b_1 x + b_2). \quad (2)$$

This parametrization allows one to include the quantities b_1 and b_2 into the optimized parameters of the articulatory model for an arbitrary speaker. For speakers that participated in the MRI experiments, the parameters b_1 and b_2 were determined as a result of minimization of the error between the function $\bar{P}(x)$ and the curve y described by Eq. (2) by the quasi-Newtonian method in the least-squares sense.

Judging from the MRI data, the shape of the cross section in the first region is close to an ellipse. It area is approximated as

$$S_1 = (\pi/4) r_1 r_2, \quad (3)$$

where $r_1(x)$ is the measured distance in the saggital plane and $r_2(x)$ is the measured width of the tract.

From Figs. 3a, 3b, 4a, and 4b, one can see that, in the second and third regions, the shape of the back wall of the tract, which is represented by the upper half of the section, is close to parabolic. Then, for a flat tongue surface, the cross-section area is calculated as

$$S_2 = S_3 = \frac{2}{3} r_3 r_4, \quad (4)$$

where $r_3(x)$ is the measured distance from the tongue surface to the rear surface of the tract and $r_4(x)$ is the width of the parabola basement, which is equal to the measured width of the tract.

In the fourth region, the parabolic shape of the rear wall of the tract is distorted by the shape of the uvula, which, in its turn, can be approximated by another parabola (Figs. 3c and 4c). Then, under the same assumption that the tongue surface is flat, the cross-section area can be calculated as

$$S_4 = \frac{2}{3} r_5 r_6 - \frac{2}{3} r_7 r_8, \quad (5)$$

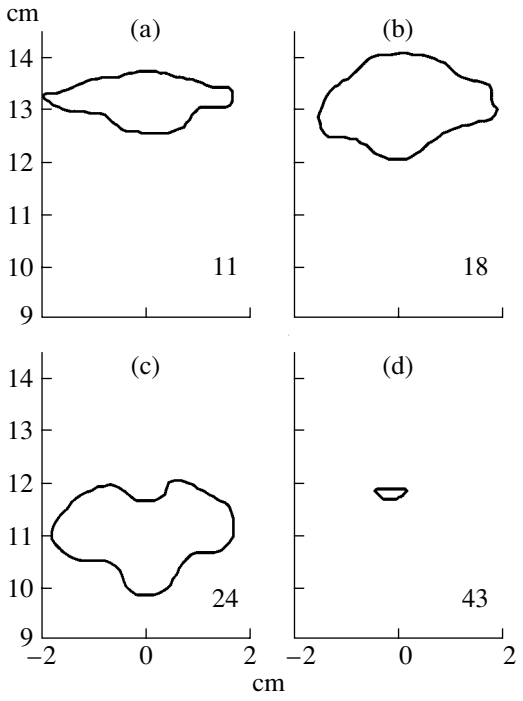


Fig. 3. Some shapes of the cross section for the vowel [i]. The numbers above the plots correspond to the numbers of plates shown in Fig. 1.

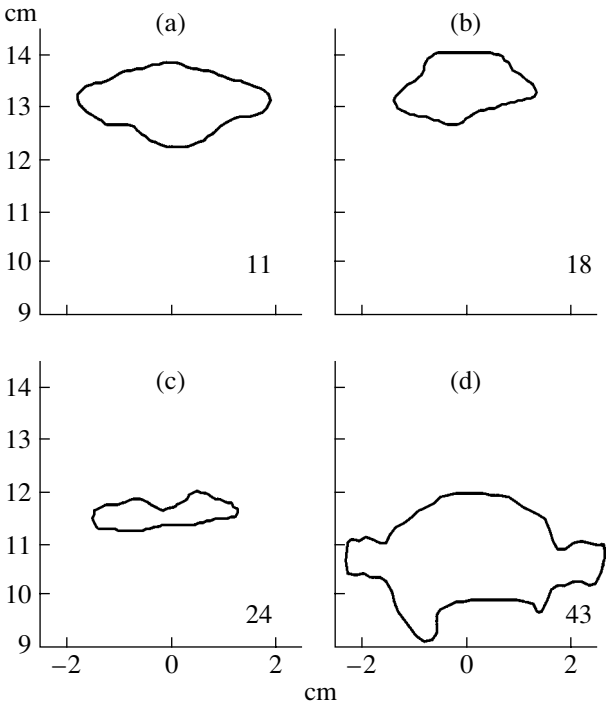


Fig. 4. Some shapes of the cross sections for vowel the [u] with the same speaker as in Fig. 3. The numbers above the plots correspond to the numbers of plates shown in Fig. 1.

where $r_5(x)$ is the measured width of the tract, $r_6(x)$ is the measured distance from the tongue surface to the uvula, r_7 is the measured width of the uvula, and r_8 is the measured thickness of the uvula.

The shape of the cross section in the fifth region depends on the tongue position with respect to the upper and lower teeth. If the tongue is above the plane tangential to the teeth of the upper jaw, the cross section has the shape shown in Fig. 3d. In this case, the area of the cross section can be approximated by the area of a parabola whose height is equal to the distance from the tongue surface to the hard palate and whose base width is determined not only by the distance from the tongue surface to the upper palate but also by the distance from the lower jaw base to the coordinate of the given cross section:

$$S_5 = \frac{2}{3} W r_9 \sqrt{\left(1 - \frac{m}{c}\right)\left(\frac{r_9}{R_9}\right)}, \quad (6)$$

where W is the measured width of the base of the lower jaw, $r_9(x)$ is the measured distance from the tongue surface to the hard palate, $R_9(x)$ is the measured distance from the line of intersection of the hard-palate surface and the plane in which the cross-section area is measured, c is the measured length of the upper jaw from the base to the front surface of the cutting teeth, and $m(x)$ is the distance from the line connecting the molar teeth with the plane in which the cross-section area is measured.

If the tongue is below the plane touching the teeth of the upper jaw, the cross section has the shape shown in Fig. 4d and its area is calculated as

$$S_5 = \frac{2}{3} W R_9 \sqrt{1 - \frac{m}{c}} + 2 l t_1 + (r_9 - R_9)\sqrt{1 - \frac{m}{c}}, \quad (7)$$

where t_1 is the measured thickness of the molar teeth and l is the measured distance from the tips of the upper teeth to the tips of the lower teeth.

In the sixth region, the cross-section area depends on the height of the tongue and also on the tongue position with respect to the front lower teeth. If the tongue moves backwards and the sublingual surface is revealed, the cross-section area is calculated by Eq. (7), where t_1 is the measured thickness of the cutting teeth and $r_9(x)$ is the measured distance from the sublingual surface to the hard palate. If the tongue touches the front lower teeth, the cross-section area is also calculated by Eq. (7), but in this case $r_9(x)$ is the measured distance from the tongue surface to the hard palate.

The genioglossus muscle is attached at its one end to the lower jaw, and its other end grows into the tongue body. The activity of the rear and middle parts of this muscle may lead to an axial deflection of the tongue (Figs. 3a–3c, 4a, and 4b). The area of this deflection can be approximated as

$$S_{ax}(x, y) = a_{ax}(x) \sin\left(\frac{\pi y}{l_{ax}}\right), \quad (8)$$

Table

Vocal tract region	Mean relative error (%)		Standard deviation of relative errors (%)	
	proposed algorithm	αβ model	proposed algorithm	αβ model
1	17.3%	20.4%	7%	35.4%
2	0.04%	10.3%	10.5%	25.0%
3	–4.6%	–3.0%	9.3%	33.7%
4	5.0%	4.2%	41.3%	43.0%
5, 6	6.4%	8.4%	13.5%	40.2%

where l_{ax} is the measured width of the tongue in this cross section and $a_{ax}(x)$ is the measured amplitude of the deflection [5].

The MRI data for lips are absent. Previous X-ray studies [5] showed that the shape of each lip is approximately described by a sinusoid half-period. Hence, the area of the lip opening can be calculated as

$$S_7 = 2R_l(x)\int_0^{l_l} \sin\frac{\pi y}{l_l}dy = \frac{4R_l l_l}{\pi}, \quad (9)$$

where $R_l(x)$ is the distance between the lips and l_l is the current lip length.

The accuracy of different algorithms used for calculating the cross-section areas was investigated in [12]. It was found that the highest accuracy is achieved with the αβ model. This model was programmed and applied to our data. The parameters α and β were determined by the minimization of the following functional for each of the tract regions:

$$F(\alpha, \beta) = \sum_{i=1}^{N_K} (S_i - \alpha(h(i))^{\beta})^2, \quad (10)$$

where K is the total number of regions into which the vocal tract was divided, N_K is the number of plates in the Kth region, i is the current number of a plate, and $h(i)$ is the measured saggital distance of the tract in the ith plate.

The results of approximating the measured cross-section areas of the vocal tract are shown in the table. The relative error was assumed to be positive if the area calculated by the algorithm proved to be greater than the measured area. Otherwise, the relative error was negative. Since, for the alveolar region, the data were not obtained for all sounds, the fifth and sixth regions were combined into one.

From the table, one can see that the mean values of errors are approximately the same for the two algorithms, except for the lower pharynx region. In this region, our algorithm provides a much smaller error, thanks to the use of the active variation of the pharynx width. In addition, the difference between the algorithms manifests itself in the error variance, which for the proposed algorithm is 2–5 times smaller than for the αβ model. This testifies to the higher accuracy and stability of the proposed algorithm as compared to the αβ model. An exception is part of the velum, where the standard deviation is fairly large for both methods.

In analyzing the results shown in the table, it is necessary to take into account the evaluation of the accuracy of the MRI method. Since the spatial resolution of the scanner is 0.1 cm/pixel, the minimum measurable area is 0.010 cm^2. For example, an area of 0.2 cm^2 should be represented by 20 pixels. Therefore, if the algorithm that performs the processing of the MRI results fails in one pixel, the error in the measured area with respect to the actual cross-section area is 5%.

An algorithm for estimating the accuracy of the MRI data was proposed in [6]. Knowing the cross-section area, one can calculate the circumference of a circle whose area is equal to the measured one. The error in the area measurement is determined by the pixels lying on the circumference. If we assume approximately that every pixel lying on the circumference is separated in two halves by the circumference, the number of pixels on the circumference is determined to be half the ratio of the circumference to the spatial resolution of the plate. Then, the MRI measurement error for an area of 0.22 cm^2 is close to 40%.

In our experiments, for those regions of the vocal tract where the mean areas were about 0.2 cm^2, the mean error between the calculated and measured cross-section areas proved to be about 15–20% and the standard deviation was within 40–45%. Thus, the approximation errors for areas of an order of 0.2 cm^2 fall within the errors in the MRI measurements of the cross-section areas. The source of the large values of the standard deviation in the velum region is the considerable scatter of the area values in this region. For some vowels, a nasalization, i.e., a lowering of the velum, was observed. In this case, the mean areas in this region were of the order of 0.22 cm^2, which resulted in large values of the standard deviation.

According to [6], the MRI measurement error for an area of about 3 cm^2 is about 7%. In our experiments, for those regions of the vocal tract where the mean areas were about 3 cm^2, the mean approximation error was about 2–5% (i.e., within the MRI measurement error for cross-section areas) and the standard deviation was about 5–8%. The variance of the errors characterizing the approximation of the cross-section areas of the vocal tract by our algorithm also proves to be within the measurement error, which testifies to the adequacy of the chosen approach.

Formulas (3)–(9) approximating the shape of the cross section of the vocal tract were used to construct an algorithm for calculating the cross-section areas from the articulatory parameters of the mathematical model of articulation.

The algorithm operates as follows. At the first stage, the vocal tract is separated into seven regions, six of

which are shown in Fig. 2 and the seventh of which represented by the lips. Then, from the current values of the controlled saggital parameters of the mathematical articulation model, the function $h(x)$ is calculated [5]. The factorial loads a_i involved in Eq. (1) are included in the set of controlled parameters of our articulation model. From the current values of these parameters and also from the preset values of the parameters b_1 and b_2, the pharynx width is calculated using Eqs. (1) and (2). The eigenvectors of the pharynx width f_i are taken from the database. From the calculated pharynx width, the parameters r_2, r_4, and r_5 involved in Eqs. (3)–(5) are determined. From the function $h(x)$, the saggital distances between the moving and immovable surfaces of the tract, which are involved in Eqs. (3)–(9), are determined for different regions. This allows one to calculate the distribution of the cross-section areas in the vocal tract (from the glottis to the alveoles) by using the aforementioned formulas.

The proposed algorithm for calculating the cross-section areas from the articulatory control vector $u(t)$ and the anatomic parameters was used to solve the inverse problem of determining the shape of the vocal tract from acoustic and geometric data [14, 15]. The inverse problem was solved for vowels, vowel + vowel and vowel + fricative + vowel combinations, and some American English words pronounced by a male speaker. The input data were the first three resonant frequencies measured in the speech signal (for vowels); the measured acoustic spectra and some spectral parameters (for fricative sounds); and the corresponding trajectories of eight moving points on the inner surfaces of the vocal tract, which were determined by a microbeam X-ray study. The shape of the hard-palate vault, the dimensions of the upper and lower jaws, and the height and width of teeth were taken from the database formed on the basis of the microbeam X-ray data [13]. The parameters a_1, a_2, b_1, and b_2, along with the articulatory parameters, were included in the optimization process in solving the inverse problem. When calculating the transfer function of the vocal tract, the optimization process was performed with allowance for the compliance of the tract walls and the presence of branching at the esophagus inlet [16]. The mean error in the coordinates of the points was 2.8%; the mean error in the first resonance, 3.7%; the mean error in the second resonance. 3.8%; and the mean error in the third resonance, 2.6%. The mean error between the measured fricative spectra and those calculated in the optimization process was about 30%. The dynamic solution obtained for the inverse problem was used by an articulatory synthesizer to generate a speech signal. The synthesized speech sounds and their sonogram proved to be very close to the original sounds. The original and resynthesized speech sounds can be displayed from the sound files attached to [14, 15].

Thus, the use of the possibilities offered by MRI measurements allowed us to construct a new algorithm for calculating the cross-section areas of the vocal tract on the basis of a three-dimensional articulation model. This algorithm takes into account not only the anatomic parameters and shape of the vocal tract in the mid-saggital plane but also the articulatory parameters in the frontal and axial planes. The algorithm developed by us was successfully tested in solving the acoustic parameters–to–vocal tract shape inverse problem by varying the parameters of the articulatory model. The use of the factor of varying pharynx width has made it possible to considerably reduce the error in calculating the shape of the vocal tract and the acoustic characteristics with respect to the measured ones.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 03-01-00116).

REFERENCES

1. D. Beautemps, P. Badin, and G. Bailly, J. Acoust. Soc. Am. **109**, 2165 (2001).
2. J. Heinz and K. Stevens, J. Acoust. Soc. Am. **36** (A), 1037 (1964).
3. J. Sundberg, *On the Problem of Obtaining Area Functions from Lateral X-ray Pictures of the Vocal Tract* (STL QPSR, Stockholm, 1969), pp. 43–45.
4. M. Tiede, H. Yehia, and E. A. Vatikiotis-Bateson, in *Proceedings of 1st ESCA Tutorial and Research Workshop in Speech Production Modeling: From Control Strategies to Acoustics* (Autrans, France, 1996), p. 41.
5. V. N. Sorokin, *Speech Synthesis* (Svyaz', Moscow, 1992) [in Russian].
6. B. Story, I. Titze, and E. Hoffman, J. Acoust. Soc. Am. **100**, 537 (1996).
7. P. Badin, G. Bailly, M. Raybadi, and C. A. Segebarth, in *Proceedings of 3rd ESCA/COCOSDA International Workshop on Speech Synthesis* (1998), p. 249.
8. P. Badin and O. Engwall, in *Proceedings of 5th Seminar on Speech Production* (Kloster Seeon, Germany, 2000), p. 297.
9. I. Titze, B. Story, and E. Hoffman, J. Acoust. Soc. Am. **104**, 471 (1998).
10. F. Bell-Berti, J. Acoust. Soc. Am. **57**, 456 (1975).
11. T. Baer, J. C. Gore, L. C. Gracco, and P. W. Nye, J. Acoust. Soc. Am. **90**, 799 (1991).
12. A. Soquet, V. Lecuit, T. Metens, and D. Demolin, Speech Commun. **36**, 169 (2002).
13. J. Westbuty, *X-ray Microbeam Speech Producton Database. User's Handbook. Version 1* (1994).
14. A. S. Leonov, I. S. Makarov, V. N. Sorokin, and A. I. Tsyplikhin, Inf. Protsessy **3** (2), 71 (2003); www.jip.ru.
15. A. S. Leonov, I. S. Makarov, V. N. Sorokin, and A. I. Tsyplikhin, Inf. Protsessy **4** (2), 141 (2004).
16. I. S. Makarov and V. N. Sorokin, Akust. Zh. **50**, 389 (2004) [Acoust. Phys. **50**, 323 (2004)].

Translated by E. Golyamina

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 44–51. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 59–67.

Conservation Laws in Acoustics

Yu. I. Bobrovnitskiĭ

Blagonravov Institute of Mechanical Engineering, Russian Academy of Sciences,
Malyi Khariton'evskii per. 4, Moscow, 101990 Russia
e-mail: bobrovni@orc.ru

Received August 31, 2004

Abstract—The method earlier proposed by the author for obtaining a complete set of conservation laws is applied to a number of simple acoustic problems. Some of the conservation laws presented are derived for the first time. Special attention is given to the physical interpretation of the results. © *2005 Pleiades Publishing, Inc.*

The theory of invariants is a large and much investigated field of science, which plays a key role in some areas of mathematics and physics [1]. In acoustics and theory of oscillations, the invariance of the basic equations with respect to different groups of transformations, for example, canonical ones, as well as the general invariants, including the volume and circulation of trajectories in the state space (the Liouville and Helmholtz theorems), Poincare integral invariants, etc., are of general theoretical importance [2]. For analyzing wave and oscillatory processes in specific media and dynamic systems, the invariants of most practical importance are those called conservation laws or, otherwise, integrals of motion. These invariants include, for example, the energy-conservation law, the momentum-conservation law, the energy-flux-conservation law, etc. [2, 3]. The knowledge of the quantities that conserve their values in the course of motion often allows one to describe the physics of phenomena in sufficient detail without solving any equations. The present paper is devoted to the construction and analysis of precisely this kind of invariant.

Conservation laws are closely related to the symmetry properties of the medium or system under consideration. Jacobi [4] presumably was the first to notice this fact. Owing to the efforts of Hilbert, Klein, and Noether, this idea evolved into a general result, which now is known as the Noether theorem [5]. According to this theorem, each transformation of variables that does not change the Lagrangian of a given physical system corresponds to a specific conservation law and vice versa. The theorem allows one to represent the conservation laws in an explicit form if the symmetry transformation is represented analytically. For example, if the Lagrangian of the system does not depend on time or, in other words, the system possesses a translational symmetry in time (the corresponding symmetry transformation is a time shift), the Noether theorem yields the energy-conservation law. According to this theorem, the homogeneity and isotropy of the space lead to the momentum- and moment-of-momentum-conservation laws, and so on [2, 3]. Unfortunately, the Noether theorem is only applicable to physical systems whose motion is described by differential equations obtained from the variational principle. Therefore, numerous publications that followed Noether's paper [5] were devoted not only to specific applications of the theorem but also to generalizations of the group theoretical approach to systems for which the classical Lagrange function did not exist, for example, to damped and gyroscopic oscillatory systems [6–13]. In particular, attempts were made to modify the variational principle [8] or to construct an unconventional Lagrangian which, under a formal application of the classical variational principle and the Noether theorem, could provide a correct result [7]. In [10], the problem of free oscillations for a system with a viscous loss was first reduced through a change of variables to the problem for a similar lossless system and, then, the conservation laws known for a conservative system were transformed into the conservation laws for the initial (nonconservative) system by returning to the initial variables. Some authors obtained the integrals of motion directly from the equations of motion by using artificial procedures (see, e.g., [11–13]). However, no general method was developed for constructing a complete set of irreducible conservation laws. Note that the above brief review refers to only a small part of the extensive literature concerned with conservation laws, namely, to the part most closely related to the present paper.

In a recent publication [14], a new method was proposed for obtaining conservation laws, and this method partially solves the problem under discussion. The method can be applied to linear physical systems in which the oscillatory or wave processes are described by sets of ordinary linear differential equations, including systems to which the Noether theorem does not apply. The method allows one to construct a complete set of independent conservation laws of the bilinear (quadratic or energy) type. In the present paper, the method proposed in [14] is applied to a number of simple acoustic problems. Some of the conservation laws

1063-7710/05/5101-0044$26.00

given below are obtained for the first time. Special attention is paid to the physical interpretation of the results.

Briefly, the approach used to obtain the conservation laws is as follows [14]. Consider a linear inhomogeneous continuous medium or a discrete oscillatory system whose motion is described by a linear differential operator of the nth order:

$$l(y) = a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \ldots + a_{n-1}(x)y' + a_n(x)y, \quad (1)$$

where x is the space or time coordinate, $y(x)$ is the field quantity (displacement, pressure, etc.), and the prime and superscript (k) denote the first and kth derivatives. The equation with operator (1) can be represented in several equivalent forms and, in particular, as a set of n equations of the first order. Here, we consider only form (1), although the method is also applicable to sets of equations (see Example 3 below). Operator (1) describes a rather wide class of acoustic situations: oscillations of linear finite-element (discrete) models, waves in one-dimensional continuous media, and some wave processes in arbitrary fluid and solid waveguides and two- and three-dimensional inhomogeneous media.

The general solution to the homogeneous equation

$$l(y) = 0 \quad (2)$$

has the form

$$y(x) = \sum_{j=1}^{n} c_j y_j(x), \quad (3)$$

where $y_j(x)$ are linearly independent particular solutions to Eq. (2) and c_j are arbitrary constants. The conservation law (integral of motion) of solution (3) is a function that does not depend on x in the interval under consideration: $P[y(x)] = \text{const}(x)$. The constant on the right-hand side of the latter equality is a combination of constants c_j and parameters of the problem. The number of conservation laws is infinite, but only n of them are independent (basic). All other laws are expressed through the basic ones. The method proposed in [14] constructs n basic conservation laws of the bilinear (quadratic) type.

The derivation of the conservation laws is based on the well-known Lagrange identity [15]

$$l(y)\bar{z} - y\overline{l^*(z)} = \frac{d}{dx}P(y, z). \quad (4)$$

Here, l^* is the nth-order differential operator adjoint to l, $P(y, z)$ is the so-called bilinear form containing derivatives from y and z up to the $(n-1)$th order (expressions for l^* and P are given in [14, 15]), and the overbar means complex conjugation. Note that operators and functions involved in Eqs. (1)–(4) are assumed to be complex in the general case and Hermitian products of the $a\bar{b}$ type are accepted for them.

If $y(x)$ is a solution to Eq. (2) and $z(x)$ is a solution to the adjoint equation

$$l^*(z) = 0, \quad (5)$$

from Lagrange identity (4) it follows that the bilinear form does not depend on x and, hence, represents a conservation law (let us call it the first conservation law):

$$P_1(y, z) = P(y, z) = C_1. \quad (6)$$

If, in identity (4), we replace solution y by y', we obtain the second conservation law [14]:

$$P_2(y, z) = P(y', z) + \int_{x_0}^{x} l'(y)\bar{z}dx = C_2, \quad (7)$$

where l' represents operator (1), in which coefficients $a_k(x)$ are replaced by $a_k'(x)$ while, in the bilinear form $P(y', z)$, the term with the nth derivative is replaced by terms with lower order derivatives using Eq. (2). The third conservation law is obtained from identity (4) by replacing y with y'':

$$P_3(y, z) = P(y'', z) + \int_{x_0}^{x} [2l'(y') + l''(y)]\bar{z}dx = C_3. \quad (8)$$

Further, replacing y by $y^{(k)}$ and z by $z^{(m)}$ in Lagrange identity (4), we obtain n basic conservation laws and a number of other laws representing their combinations.

One more group of bilinear conservation laws can be obtained by proceeding from the Lagrange identity with ordinary ab-type products of complex quantities:

$$l(y)z - y\bar{l}(z) = \frac{d}{dx}P(y, z). \quad (9)$$

Repeating the derivation, we arrive at conservation laws in the form

$$P_1(y, z) = P(y, z) = C_1,$$
$$P_2(y, z) = P(y', z) + \int_{x_0}^{x} l'(y)zdx = C_2, \text{ etc.} \quad (10)$$

These laws coincide with laws (6)–(8) as the adjoint operators and bilinear forms coincide, provided that the equations and their solutions are real. In the complex case, they differ in form and in physical meaning (see examples below). Note that, while conservation law (6) is well known, laws (7), (8), and (10) have never before been mentioned in the literature (as far as the author of this paper knows). For media and systems whose oscillations are described by differential operators with constant coefficients, the conservation laws obtained above have especially simple forms, because they contain no integrals. For layered and other media that are described by differential operators (1) with nondifferentiable or discontinuous coefficients, these conservation laws remain valid if one passes from the common

derivatives to the so-called quasi-derivatives (see [14]). Now, let us consider some examples.

Example 1. To reveal the physical meaning of the conservation laws obtained above, let us consider the simplest system: a lossless oscillator (a mass m attached to a spring with a stiffness χ). Free oscillations of this system are described by an ordinary differential equation of the second order:

$$l(y) = m\ddot{y} + \chi y = 0, \tag{11}$$

where $y(t)$ is the displacement of the mass, the parameters m and χ are assumed to be real and independent of time t, and an overdot means a derivative with respect to t. Operator (11) is a self-adjoint one, i.e., $\bar{l} = l^* = l$, and its bilinear form is equal to $P(y, z) = m\dot{y}z - my\dot{z}$. The general solution to an equation of the second order depends on two arbitrary constants, and, hence, two independent conservation laws are obtained for it:

$$P_1(y, z) = P(y, z) = m\dot{y}z - my\dot{z} = C_1, \tag{12}$$

$$P_2(y, z) = P(y, \dot{z}) = m\dot{y}\dot{z} + \chi yz = C_2. \tag{13}$$

All other conservation laws are combinations of laws (12) and (13). For example, the law

$$P_3(y, z) = P(\dot{y}, \dot{z}) = -\chi y\dot{z} + \chi\dot{y}z = C_3 \tag{14}$$

is equivalent to law (12) with the factor χ/m, and so on. Functions $y(t)$ and $z(t)$ involved in Eqs. (12)–(14) represent two solutions to Eq. (11). Let us first consider the case when these solutions are real functions of time and the Hermitian product coincides with the ordinary one.

Let

$$y(t) = A_1\cos\omega_0 t + A_2\sin\omega_0 t$$

be the general solution to Eq. (11), so that this solution depends on two real arbitrary constants A_1 and A_2 and $\omega_0 = (\chi/m)^{1/2}$ is the eigenfrequency. First, let us show that, for any solution $z(t)$, the constants on the right-hand sides of conservation laws (12)–(14) are combinations of constants A_1 and A_2 and the oscillator parameters. Indeed, for $z_1(t) = \cos\omega_0 t$, from Eqs. (12) and (13), we obtain

$$P_1(y, z_1) = C_1 = 2m\omega_0 A_2,$$
$$P_2(y, z_1) = C_2 = 2\chi A_1.$$

For the other independent solution to Eq. (11), $z_2(t) = \sin\omega_0 t$, laws (12) and (13) yield

$$P_1(y, z_2) = C_1 = -2m\omega_0 A_1,$$
$$P_2(y, z_2) = C_2 = 2\chi A_2.$$

If solution $z(t)$ coincides with $y(t)$, law (12) becomes an identity ($C_1 \equiv 0$) and carries no useful information, while law (13) represents the conservation law for the total instantaneous energy $E(t)$ of the system:

$$P_2(y, y) = m\dot{y}^2 + \chi y^2 = 2E = C_2 = 2\chi(A_1^2 + A_2^2). \tag{15}$$

This result is well-known: the energy of free oscillations of a conservative system does not depend on time. Note that law (15) is the only conservation law that is encountered in the literature for a lossless oscillator. Law (13) obtained above is a more general conservation law, because it relates two different free motions of the system; it coincides with law (15) at $z = y$. Conservation law (12), which also relates two different solutions to Eq. (11), has never been encountered by us in the literature.

To reveal the physical meaning of conservation laws (12)–(14), it is necessary to introduce the notation of cross-energy quantities. For two oscillatory motions, $y(t)$ and $z(t)$, the quantity $T_{yz}(t) = m\dot{y}\dot{z}/2$ can be called the kinetic cross-energy of the momentum $m\dot{y}$ of the first motion at the velocity $\dot{z}$ of the second motion, and the quantity $T_{zy}(t) = m\dot{z}\dot{y}/2$ can be called the instantaneous kinetic cross-energy of the momentum of the second motion at the velocity of the first motion: $T_{yz}(t) = T_{zy}(t)$. In a similar way, we introduce the potential cross-energy of the two motions:

$$U_{yz}(t) = U_{zy}(t) = \frac{1}{2}\chi yz,$$

as well as the total cross-energy and the cross-Lagrangian: $E_{yz} = T_{yz} + U_{yz}$, …, $L_{zy} = T_{zy} - U_{zy}$. These quantities naturally appear when quadratic (energy) quantities of the sum of motions are considered. For example, the total energy of the sum of two motions $y + z$ is equal to the sum of energies of individual components and the cross-energies:

$$E = E_{yy} + E_{zz} + E_{yz} + E_{zy}.$$

Returning to conservation law (13), one can easily see that, physically, this law means that, for any two free motions of a lossless oscillator, the following quantities are conserved (do not depend on time):
the cross-energies

$$2E_{yz}(t) = 2E_{zy}(t) = m\dot{y}\dot{z} + \chi yz = C_2 \tag{16}$$

and the energies of individual motions

$$2E_{yy}(t) = m\dot{y}^2 + \chi y^2 = C_2',$$
$$2E_{zz}(t) = m\dot{z}^2 + \chi z^2 = C_2''. \tag{17}$$

Let us also introduce the cross-action of the two motions of the oscillator: after some transformation, this quantity can be reduced to the form

$$S_{yz}(t) = \int_{t_0}^{t} L_{yz}\,dt = \frac{1}{2}m\dot{y}z - S_{yz}(t_0).$$

From this expression, one can see that conservation law (12) can be physically interpreted as follows: for any two free motions of the oscillator, the action of the momentum of the first motion $m\dot{y}$ on the displacement

z of the second motion varies in time in the same way as the action of the momentum $m\dot{z}$ of the second motion on the displacement of the first motion but differs from it by a time-independent constant,

$$S_{yz}(t) - S_{zy}(t) = C_1. \tag{18}$$

Although other conservation laws are combinations of laws (12) and (13), they may have other physical meaning. In particular, since the first term in law (14) represents the cross-power

$$F_{yz}(t) = (-\chi y)\dot{z},$$

i.e., the power of the force $(-\chi y)$ acting on the mass in the first motion and on the velocity $\dot{z}$ of the mass in the second motion, conservation law (14) can be interpreted as the constancy of the difference between two cross-power fluxes,

$$F_{yz}(t) - F_{zy}(t) = C_3. \tag{19}$$

One more useful aspect of the physical interpretation of the conservation laws obtained above is related to the behavior of energy quantities as functions of time. Let $y(t)$ and $z(t)$ be arbitrary free motions of the oscillator:

$$\begin{aligned} y(t) &= A_1\cos\omega_0 t + A_2\sin\omega_0 t, \\ z(t) &= a_1\cos\omega_0 t + a_2\sin\omega_0 t. \end{aligned} \tag{20}$$

Any quadratic or bilinear functions of signals that are harmonic in time, including the energy quantities involved in the conservation laws, consist of two components: a constant component (i.e., the time-averaged component) and a variable component varying in time with a double frequency [16]. For example, the kinetic cross-energy $T_{yz}(t)$ for motions (20) is equal to the sum of the constant and variable components $T^{=}_{yz}$ and $T^{\approx}_{yz}(t)$, where the latter, in its turn, consists of the cosine T^{c}_{yz} and sine T^{s}_{yz} components:

$$\begin{aligned} T_{yz}(t) &= T_{zy}(t) = T^{=}_{yz} + T^{\approx}_{yz}(t), \\ T^{=}_{yz} &= \frac{1}{4}m\omega_0^2(A_1a_1 + A_2a_2), \\ T^{\approx}_{yz}(t) &= T^{c}_{yz}\cos 2\omega_0 t + T^{s}_{yz}\sin 2\omega_0 t, \\ T^{c}_{yz} &= \frac{1}{4}m\omega_0^2(-A_1a_1 + A_2a_2), \\ T^{s}_{yz} &= -\frac{1}{4}m\omega_0^2(A_1a_2 + A_2a_1). \end{aligned} \tag{21}$$

The potential cross-energy of the system for motions (20) also consists of the constant and variable components, where the first of them is equal to the constant component of kinetic energy $U^{=}_{yz} = U^{=}_{zy} = T^{=}_{yz} = T^{=}_{zy}$ (this is a general law: the kinetic energy of free oscillations of a lossless system is, on average, equal to the potential energy [17]), and the second differs in sign from the variable component of kinetic energy (21). As a result, the total cross-energy only contains a nonzero constant component, so that conservation law (16) is equivalent to the absence of variable component in the instantaneous energy:

$$\begin{aligned} E^{\approx}_{yz}(t) &= E^{\approx}_{zy}(t) = 0, \\ E_{yz}(t) &= E_{zy}(t) = E^{=}_{yz} = \text{const}. \end{aligned} \tag{22}$$

The same is true for laws (17).

One can directly verify that the cross-power fluxes $F_{yz}(t)$ and $F_{zy}(t)$ have constant components that are equal in magnitude but opposite in sign and variable components that are identical to each other:

$$F^{=}_{yz} + F^{=}_{zy} = 0, \quad F^{\approx}_{yz}(t) - F^{\approx}_{zy}(t) = 0. \tag{23}$$

Similar relations are valid for action. Therefore, conservation laws (18) and (19) can be interpreted as the equality of the variable components of cross-powers and actions and, hence, as the independence of the difference of their instantaneous values from time.

Now, let us consider the case of complex solutions to Eq. (11). Complex solutions are often used in acoustics. Although in considerations their real parts are always implied, these solutions have their own independent meaning. Therefore, we consider more closely the physical meaning of the conservation laws mentioned above in the case of complex solutions. This meaning essentially depends on the type of the product of the complex functions. If an ordinary product of two complex functions, e.g., YZ, is used with initial Lagrange identity (9), the conservation laws for an oscillator exactly coincide with laws (12)–(14). If a Hermitian product $Y\bar{Z}$ is used with initial identity (4), the conservation laws can be represented as

$$P_1(Y, Z) = m\dot{Y}\bar{Z} - mY\dot{\bar{Z}} = C_1, \tag{24}$$

$$P_2(Y, Z) = m\dot{Y}\dot{\bar{Z}} + \chi Y\bar{Z} = C_2, \tag{25}$$

$$P_3(Y, Z) = -\chi Y\dot{\bar{Z}} + \chi\dot{Y}\bar{Z} = C_3. \tag{26}$$

Let

$$Y(t) = A\exp(-i\omega t), \quad Z(t) = a\exp(-i\omega t) \tag{27}$$

be two arbitrary complex solutions to Eq. (11) with complex amplitudes A and a, respectively. Their real parts coincide with solutions (20), provided that

$$A = A_1 + iA_2, \quad a = a_1 + ia_2. \tag{28}$$

Consider the kinetic cross-energy of complex oscillations (27) for an ordinary product of complex quantities:

$$\tilde{\tilde{T}}_{YZ}(t) = \frac{1}{2}m\dot{Y}\dot{Z} = -\frac{1}{2}m\omega_0^2 Aa\exp(-i2\omega_0 t). \tag{29}$$

One can see that this is a complex harmonic function of double frequency without any constant component. Substituting Eqs. (28) into Eq. (29), one can verify that the real part of function (29) is twice the variable component of instantaneous kinetic energy (21): $\mathrm{Re}(\tilde{\tilde{T}}_{YZ}(t)) = 2\tilde{\tilde{T}}_{yz}(t)$. As for the imaginary part of function (29), which can be called the reactive variable component of the kinetic cross-energy, it does not have such a direct physical meaning. One can show that the imaginary part is a Hilbert transform of the physically significant real variable component and, hence, complements it to form the analytical (complex) function of time given by Eq. (29) (see, e.g., [18]). However, when operations are performed with physical quantities in the complex domain, the reactive (imaginary) parts play a role that is equally important as that of the active (real) parts. For example, the reactive part of function (29) multiplied by $-2\omega_0$ characterizes the variation (time derivative) of the active part of the kinetic cross-energy and, hence, is equal to the active part of the cross-power flux $\mathrm{Re}(\tilde{\tilde{F}}_{YZ})$, and so on.

The potential cross-energy of complex solutions (27) for the ordinary product

$$\tilde{\tilde{U}}_{YZ}(t) = \frac{1}{2}\chi YZ = \frac{1}{2}\chi Aa\exp(-i2\omega_0 t)$$

also describes only the variable component of the instantaneous energy and differs from the kinetic cross-energy in sign. Hence, the total cross-energy of solutions (27) is equal to zero:

$$\tilde{\tilde{E}}_{YZ}(t) = \tilde{\tilde{T}}_{YZ}(t) + \tilde{\tilde{U}}_{YZ}(t) = 0, \qquad (30)$$

and this is also true for the cross-energy $\tilde{\tilde{E}}_{ZY}(t)$ and for the individual energies: $\tilde{\tilde{E}}_{YY}(t) = \tilde{\tilde{E}}_{ZZ}(t) = 0$. Thus, the meaning of conservation law (13), (30) for complex solutions (27) with an ordinary product consists in that the variable components of total cross-energies and individual energies are equal to zero (see Eqs. (22)). As for conservation laws (12) and (14), these laws for the complex solutions with an ordinary product mean that the variable components of the cross-powers and actions are identical (see Eqs. (23)).

Now, let us assume that a Hermitian product of complex functions is used. It is evident that any bilinear or quadratic function of complex harmonic motions of type (27) does not depend on time and, hence, is an integral of motion. For example, in this case the kinetic cross-energy has the form

$$\bar{\bar{T}}_{YZ} = \frac{1}{2}m\dot{Y}\dot{\bar{Z}} = \frac{1}{2}m\omega_0^2 A\bar{a}. \qquad (31)$$

Since this is a complex quantity, it actually contains two independent conservation laws: for the real part and for the imaginary one. By direct calculations, one can verify that the real part of energy (31) represents twice the constant component $\bar{T}_{yz}$ of the instantaneous kinetic energy $T_{yz}(t)$ and that the imaginary part (the reactive kinetic cross-energy) is proportional to the constant component of the instantaneous power flux $\bar{F}_{yz}$. Thus, in the case of complex solutions (27) with a Hermitian product, among all conservation laws (24)–(26), only one law, e.g., law (25) or (31), is independent. All other laws can be represented as its functions; for example

$$\bar{\bar{F}}_{YZ} = -2i\omega_0\bar{\bar{T}}_{YZ} = \text{const},$$

and so on.

Let us summarize the results of considering the physical meaning of the conservation laws obtained for a lossless oscillator. Two independent laws obtained for this system correspond to the absence of time dependence for the total cross-energy and individual energy of free oscillations and to the constancy of the difference between two instantaneous cross-power fluxes or cross-actions. Since each of these quantities consists of constant and variable components, these laws are equivalent to the absence of variable component in the total energy of oscillations (i.e., the variable components of the kinetic and potential energies cancel each other) and to the equality of the variable components of two cross-power fluxes or actions.

If the natural oscillations of the oscillator are described by complex functions of time with an ordinary or Hermitian product, only one complex conservation law is independent and its real and imaginary parts are equivalent to two conservation laws in the real domain. Although the specific physical content of conservation laws depends on the system under consideration, general approaches to their interpretation, which are described above in detail for the simplest system (oscillator), are valid for other systems. They will be used below to study more complicated examples.

Example 2. Now, let us consider an oscillator with a viscous damping. Its oscillations are described by a non-self-adjoint equation

$$l(y) = m\ddot{y} + b\dot{y} + \chi y = 0 \qquad (32)$$

with real coefficients. The adjoint operator l^* differs from the direct one in the sign of the damping factor b and, thus, describes an active system with a negative loss. If the ordinary product is accepted, the bilinear form is expressed as $P(y, z) = m\dot{y}z - my\dot{z} + byz$. If the Hermitian product is accepted, z should be replaced by $\bar{z}$. For real solutions $y(t)$ and $z(t)$ to direct (32) and adjoint (5) equations, conservation laws similar to laws (12)–(14) are valid:

$$P_1(y, z) = P(y, z) = m\dot{y}z - my\dot{z} + byz = C_1, \quad (33)$$

$$P_2(y, z) = P(y, \dot{z}) = -P(\dot{y}, z) = m\dot{y}\dot{z} + \chi yz = C_2, \quad (34)$$

$$P_3(y, z) = P(\dot{y}, \dot{z}) = -\chi y\dot{z} + \chi z\dot{y} - by\dot{z} = C_3. \quad (35)$$

They are related by the formula

$$\chi P_1 - bP_2 - mP_3 = 0.$$

Therefore, only two laws are independent. As in the case of a lossless oscillator, their physical meaning consists in the conservation of some cross-energy quantities. However, unlike lossless systems, no quadratic laws are possible for lossy systems. The laws must be bilinear, and each of them should contain one solution to the direct problem and one solution to the conjugate problem. This is a consequence of the non-self-adjointness of operator (32): natural oscillations of the damped system attenuate with time, while oscillations of the adjoint system grow up. In a bilinear function, these opposite dependences compensate for each other making possible the existence of conservation laws of this type.

Law (34) represents the conservation law for the total cross-energy. It has been derived in different ways by many authors (see, e.g., [7, 9]). By analogy with law (14), law (35) can be interpreted as the equality of the variable components of two cross-power fluxes if one accepts that the force acting on the mass in the course of motion $y(t)$ is equal to $-\chi y - b\dot{y}/2$ and if expression (35) can be represented in the form

$$P_3 = (-\chi y - b\dot{y}/2)\dot{z} - (-\chi z + b\dot{z}/2)\dot{y}.$$

Analogously, conservation law (33) means that the difference between two cross-actions is independent of time if the action S_{yz} is understood as the quantity $(m\dot{y} + by/2)z$ and the action S_{zy} is equal to $(m\dot{z} - bz/2)y$. Laws (33) and (35) were never encountered by the author of this paper in the literature.

Example 3. Let us now consider oscillations described by a set of differential equations. Let these oscillations be those of a linear damped mechanical system with N degrees of freedom. They are known to be described by N linear ordinary differential equations of the second order:

$$l(y) = M\ddot{y} + B\dot{y} + Ky = 0, \quad (36)$$

where $y = [y_1, \ldots, y_n]^T$ is the function vector of displacements; M, B, and K are square matrices of order N, not necessarily real and symmetric; and the superscript T means transposition. Operator (36) with these parameters describes the oscillations of a wide class of mechanical systems with allowance for viscous and structural losses, gyroscopic effects, and Coriolis and Lorentz forces, as well as systems containing active elements. Using the Hermitian product of the function vectors $z^*y = y_1\bar{z}_1 + \ldots + y_N\bar{z}_N$, from the Lagrange identity we easily obtain the adjoint operator and the bilinear form:

$$l^*(z) = M^*\ddot{z} - B^*\dot{z} + K^*z,$$
$$P(y, z) = z^*M\dot{y} - \dot{z}^*My + z^*By.$$

The corresponding conservation laws have the form

$$P_1 = P(y, z) = C_1,$$
$$P_2 = P(y, \dot{z}) = \dot{z}^*M\dot{y} + z^*Ky = C_2. \quad (37)$$

These laws are valid for any solutions to the direct and adjoint sets of equations. The general solution to direct set of equations (36) involves $2N$ arbitrary constants. Hence, among the infinite number of laws like Eq. (37), $2N$ are the basic ones. They are obtained from Eqs. (37) with the use of some kind of N independent particular solutions to the adjoint set of equations. The physical meaning of conservation laws (37) is the same as that for a system with one degree of freedom.

Example 4. Consider the propagation of free longitudinal waves in a thin homogeneous linearly elastic rod without losses. These waves are described by the Bernoulli wave equation

$$ESu''(x, t) - m\ddot{u}(x, t) = 0, \quad (38)$$

where $u(x, t)$ is the longitudinal displacement, $m = \rho S$ is the mass per unit length, ρ and E are the density and Young's modulus of the material, and S is the cross-sectional area of the rod. All parameters are assumed to be real and independent of the spatial coordinate and time. The wave motion depends on two variables. The problem is to find the quantities that are invariable along the x axis and the quantities that are independent of time. Quantities that are simultaneously invariant with respect to both coordinates are not considered here: the generalization of the proposed method to the multidimensional case is planned to be the subject of a special publication.

Seeking the solution to Eq. (38) in the form $u(x, t) = X(x)T(t)$ by the Fourier method, for each factor we obtain the following ordinary differential equations:

$$ESX''(x) + \gamma X(x) = 0, \quad m\ddot{T}(t) + \gamma T(t) = 0, \quad (39)$$

where γ is a constant. For a motion that is harmonic (exponential) in x and t and characterized by a frequency ω and a wave number k, this constant is calculated as

$$\gamma = ESk^2 = m\omega^2, \quad (40)$$

and the general real solution to the equation under study has the form

$$u(x, t) = A_0\cos(kx - \omega t - \alpha) + B_0\cos(kx + \omega t - \beta). \quad (41)$$

This solution consists of a wave propagating with an amplitude A_0 and an arbitrary phase α in the positive direction of the x axis and a wave propagating in the opposite direction with an amplitude B_0 and phase β. The propagation velocity is $c = \omega/k = (E/\rho)^{1/2}$. In the complex domain, the solution has the form

$$u(x, t) = A\exp(ikx - i\omega t) + B\exp(-ikx - i\omega t), \quad (42)$$

where the complex amplitudes are

$$A = A_0\exp(-i\alpha), \quad B = B_0\exp(i\beta). \quad (43)$$

The real part of the complex solution given by Eqs. (42) and (43) coincides with real solution (41).

The operators involved in Eq. (39) are self-adjoint. They are analogous to operator (11) for a lossless oscillator. Hence, the conservation laws for them are analogous to laws (12) and (13). For an ordinary product, they have the form

$$P_1^x = ES(u'v - uv') = \text{const}(x), \quad (44)$$

$$P_2^x = ESu'v' + \gamma uv = \text{const}(x), \quad (45)$$

$$P_1^t = m(\dot{u}v - u\dot{v}) = \text{const}(t), \quad (46)$$

$$P_2^t = m\dot{u}\dot{v} + \gamma uv = \text{const}(t). \quad (47)$$

The superscripts indicate the coordinate along which the given quantity is conserved. These four conservation laws are valid for both real and complex solutions. For complex solutions with a Hermitian product, the conservation laws coincide with laws (44)–(47), in which the adjoint solution v is replaced by the complex conjugate one, $\bar{v}$.

The physical meaning of the quantities involved in Eqs. (44)–(47) is as follows. The quantity $f = -ESu'$ is the force acting in the cross section, and the product fv in Eq. (44) is the cross-work of this force at the displacement v. Expressions in laws (45) and (47) with allowance for Eq. (40) represent the total cross-energy density. The quantities in law (46) represent the mechanical action. For example, the term $m\dot{u}v$ is the action of the pulse $m\dot{u}$ on the displacement v (see Example 1).

For an arbitrary real solution (41) and a similar solution to the adjoint problem

$$v(x, t) = a_0\cos(kx - \omega t - \varphi) + b_0\cos(kx + \omega t - \psi),$$

conservation laws (44)–(47) are transformed into the form

$$\begin{aligned}
P_1^x/ESk &= A_0a_0\sin(\alpha - \varphi) + B_0b_0\sin(\beta - \psi)\\
&+ A_0b_0\sin(\alpha - \psi + 2\omega t) + B_0a_0\sin(\beta - \varphi - 2\omega t),\\
P_2^x/m\omega^2 &= A_0a_0\cos(\alpha - \varphi) + B_0b_0\cos(\beta - \psi)\\
&+ A_0b_0\cos(\alpha - \psi + 2\omega t) + B_0a_0\cos(\beta - \varphi - 2\omega t),\\
P_1^t/m\omega &= -A_0a_0\sin(\alpha - \varphi) + B_0b_0\sin(\beta - \psi)\\
&+ A_0b_0\sin(2kx - \alpha - \psi) - B_0a_0\sin(2kx - \beta - \varphi),\\
P_2^t/m\omega^2 &= A_0a_0\cos(\alpha - \varphi) + B_0b_0\cos(\beta - \psi)\\
&+ A_0b_0\cos(2kx - \alpha - \psi) + B_0a_0\cos(2kx - \beta - \varphi).
\end{aligned} \quad (48)$$

As one can see from Eqs. (48), the first two quantities do not depend on the spatial coordinate and contain constant components and components varying in time with a double frequency. The last two quantities in Eqs. (48) do not depend on time and contain constant components and components varying with x. It should be stressed that conservation laws (44)–(48) contain no quantities simultaneously depending on both variables. However, such quantities can be obtained by averaging Eqs. (48) over time and/or over the spatial coordinate.

When complex solutions (42), (43) with a Hermitian or ordinary product are used, the real parts of conservation laws (44)–(47) yield the constant and variable components of laws (48). In this case, the imaginary parts of some of the laws prove to be proportional to the real parts of other laws, so that not all of the laws are independent. For example, for complex solutions with a Hermitian product, the relation $\omega P_1^t + iP_2^t = 0$ is valid (see also Example 1).

It should be noted that the cross-power flux through the cross section $F_{uv} = -ESu'\dot{v}$ is not a conserved quantity, because, in addition to the constant component, it contains components varying with x and t. The sum and the difference of the cross-energy fluxes, $F_{uv} \pm F_{vu}$, also represent non conserved quantities.

Example 5. As an example of conservation laws in an inhomogeneous medium, let us consider waves in a pipe with a varying cross section (a horn). The pipe is assumed to be thin, and the variation of the cross-sectional area $S(x)$, sufficiently smooth. Then, a harmonic wave motion is approximately described by the equation [19]

$$l(p) = (Sp')' + k_0^2Sp = 0, \quad (49)$$

where $p(x)$ is the pressure in the cross section x, $k_0 = \omega/c$, and c is the velocity of sound in the medium. The operator on the left-hand side of Eq. (49) is self-adjoint: $l^* = l$. For it, independent conservation laws (6) and (7) take the form

$$P_1(p, q) = P(p, q) = S(p'q - pq') = C_1, \quad (50)$$

$$\begin{aligned}
P_2(p, q) &= P(p, q')\\
&= (Sp)'q' + k_0^2Spq - \int_{x_0}^{x} pl'(q)dx = C_2,
\end{aligned} \quad (51)$$

where p and q are any two solutions to Eq. (49). The physical meaning of the quantities involved in Eqs. (50) and (51) is the same as that in the previous example, correct to constant factors. Since, by virtue of the Euler equation [19], the derivative of pressure with respect to x is proportional to the particle velocity in the medium, conservation law (50) means that, in the cross-power fluxes, the components varying with x are identical. Conservation law (51) is related to the energy of the medium, although the meaning of its individual terms is unclear.

Let us consider some specific cases of laws (50) and (51). Let the pipe be an exponential horn, whose cross section increases according to the exponential law

$$S(x) = S_0 \exp(2\varepsilon x).$$

In this case, the general solution to Eq. (49) has the form

$$p(x) = A\exp(-\varepsilon x)\cos(kx - \alpha), \qquad (52)$$

where A and α are arbitrary constants and $k = \sqrt{k_0^2 - \varepsilon^2}$. For an exponential horn, relation $l'(q) = 2\varepsilon l(q)$ is valid and, hence, the integral in law (51) turns out to be equal to zero. If $q(x)$ is a solution similar to solution (52) with arbitrary constants a and β, the constants involved in conservation laws (50) and (51) can be represented as

$$P_1 = C_1 = kS_0 Aa\sin(\alpha - \beta),$$

$$P_2 = C_2 = k^2 S_0 Aa\cos(\alpha - \beta).$$

Solutions to Eq. (49) in terms of known functions also exist for some other horns, for example, for a power-law horn with the cross section

$$S(x) = S_0 x^n.$$

In the case of an even $n = 2(m + 1)$, the general solution to Eq. (49) can be represented in the form

$$p(x) = Ax^{-m} j_m(k_0 x) + Bx^{-m} n_m(k_0 x),$$

where j_m and n_m are Bessel and Neumann spherical functions. In the case of an odd $n = 2m + 1$, the spherical functions involved in this solution should be replaced by common cylindrical functions J_m and N_m. One can directly verify that conservation laws (50) and (51) are valid in both cases. The integral terms in Eq. (51) are not equal to zero in this case.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the President of the Russian Federation (grant no. NSh-2247.2003.1)

REFERENCES

1. E. Spencer, *Continuum Physics*, Vol. 1, Part 3: *Theory of Invariants* (McGraw-Hill, New York, 1969; Mir, Moscow, 1974).
2. H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, Mass., 1980; Nauka, Moscow, 1975).
3. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics* (Fizmatlit, Moscow, 1958; Pergamon Press, Oxford, 1965).
4. C. G. J. Jacobi, *Vorlesungen uber Dynamik* (Reimer, Berlin, 1884).
5. E. Noether, Nachr. Konig. Ges. Wiss. Goettingen, Math. **2**, 235 (1918).
6. E. Bessel-Hagen, Math. Ann. **84**, 258 (1921).
7. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953; Inostrannaya Literatura, Moscow, 1958), Vol. 1, Chap. 3.
8. D. S. Djukic and B. D. Vujanovic, Acta Mech. **23** (1–2), 17 (1975).
9. L. Y. Bahar and H. G. Kwatny, Am. J. Phys. **49**, 1062 (1981).
10. L. Y. Bahar and H. G. Kwatny, J. Sound Vibr. **102**, 551 (1985).
11. T. Honein, N. Chien, and G. Herrmann, Phys. Lett. A **155** (4), 223 (1991).
12. A. G. Kotousov and N. A. Makhutov, Dokl. Akad. Nauk **351**, 476 (1996) [Phys. Dokl. **41**, 602 (1996)].
13. U. Nordbrock and R. Kienzler, J. Sound Vibr. **256**, 981 (2003).
14. Yu. I. Bobrovnitskiĭ, Dokl. Akad. Nauk **395**, 43 (2004) [Dokl. Phys. **49**, 171 (2004)].
15. M. A. Naĭmark, *Linear Differential Operators*, 2nd ed. (Nauka, Moscow, 1969; Ungar, New York, 1967).
16. Yu. I. Bobrovnitskiĭ, Akust. Zh. **42**, 267 (1996) [Acoust. Phys. **42**, 234 (1996)].
17. G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974; Mir, Moscow, 1977).
18. E. Skudrzyk, *The Foundations of Acoustics. Basic Mathematics and Basic Acoustics* (Springer, New York, 1971; Mir, Moscow, 1976), Vol. 1.
19. P. M. Morse, *Vibration and Sound* (McGraw-Hill, New York, 1936; GITTL, Moscow, 1949).

Translated by E. Golyamina

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 5–10. From Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 9–23.

Congratulations to Akusticheskii Zhurnal on the Occasion of Its 50th Anniversary

Over the years, the journal Akusticheskii Zhurnal has been widely read and used in the United States and in other English-speaking countries. Its importance was recognized from the very beginning, and the American Institute of Physics, in response to stimulus from the Acoustical Society of America, published a translation for many years under the title "Soviet Physics–Acoustics." Since 1993, the English translation has been distributed by the American Institute of Physics under the title "Acoustical Physics." I personally have found the papers in Akusticheskii Zhurnal very fundamental and very useful in my own research. The names of many leading Russian and other Eastern European acousticians have become well known to me long before I had the opportunity to meet them in person. One may also note that many articles from Akusticheskii Zhurnal are cited in a significant manner in the textbook that I wrote in 1981. Times continue to change, but I believe that quality journals such as Akusticheskii Zhurnal will continue indefinitely.

Again, congratulations, best wishes, and with full confidence in a long and significant future.

A. D. Pierce
Editor-in-Chief
Acoustical Society of America

Vibrations of Spherical Inclusions in Elastic Solids[1]

A. D. Pierce

Department of Aerospace and Mechanical Engineering, Boston University, Boston, Massachusetts, USA
e-mail: adp@bu.edu

Received September 9, 2004

Abstract—Variational principles are derived for the analysis of dynamical phenomena associated with spherical inclusions embedded in homogeneous isotropic elastic solids. The starting point is Hamilton's principle, with the material properties assumed to vary only with the radial distance r from the origin. Attention is restricted to disturbances that are symmetric about the polar (z) axis, such that the nonzero displacement components in spherical coordinates, u_r and u_θ, are independent of the polar coordinate ϕ. The symmetry allows for a decoupling of the polar components, the nth of which is described by $U_{r,n}(r, t)P_n(\cos\theta)$ and $U_{\theta,n}(r, t)dP_n/d\theta$. A variational principle is subsequently derived for the field quantities $U_{r,n}$ and $U_{\theta,n}$. Concepts analogous to those of the theory of matched asymptotic expansions are used to embellish the principle in order to allow for the damping associated with the outward radiation of elastic waves. Examples illustrating the use of the variational principle for formulating plausible lumped-parameter models are given for the cases of $n = 0$ and $n = 1$.

1. INTRODUCTION

The present paper is concerned with a general class of fundamental dynamical problems associated with isolated inhomogeneities or foreign substances in elastic media. The general topic dates back to the analysis by Isakovich in 1949 of the oscillations of a gas-filled bubble in a nearly incompressible elastic solid. Isakovich's analysis is presented in the form of an exercise in the text on elasticity [1] in the series by Landau and Lifshitz. A more elaborate derivation for a somewhat more general case was given in 1958 by Meyer [2] and others.

A related topic is the scattering of elastic waves by inclusions of limited size. Scattering by spherical cavities was treated in an early paper by Sivukhin [3], which appeared in the first issue of *Acoustics Journal* in 1955. Ying and Truell [4] subsequently gave a lengthy analysis of compressional wave scattering from general classes of spherical inclusions, in which they made use of analytical techniques used much earlier by Herzfeld [5] for analysis of the scattering of sound waves by small elastic spheres in a fluid with viscosity.

[1] This article was submitted by the author in English.

1063-7710/05/5101-0005$26.00 © 2005 Pleiades Publishing, Inc.

Over the years, many other papers have appeared on this general topic of elastic wave scattering by spherical inclusions. Examples include the papers by Knopoff [6]; Einspruch and Truell [7]; Johnson and Truell [8]; Mow [9], Norwood and Miklowitz [10]; and Gaunaurd and Überall [11].

For the most part, the papers concerned with scattering have focused on far-field scattering and on numerical predictions of quantities associated with scattering rather than with the details of what is happening within and near the inclusion itself. A more recent trend, somewhat in the tradition of the work by Isakovich and by Meyer and his colleagues, is evident in a 1999 paper in *Acoustical Physics* by Alekseev and Rybak [12] that is concerned with the oscillations of a gas bubble encased in an elastic shell and within an external elastic medium. Later papers by Allen and Roy [13]; by Khismatullin and Nadim [14]; by Zaslavskiĭ [15]; and by Emelianov, Hamilton, Ilinskii, and Zabolotskaya [16] have examined such bubble oscillations taking into account the viscoelasticity of the medium, various possibilities regarding the shell enclosing the bubble, and nonlinear effects.

The present paper recognizes that vibrations that are not purely radial may be intrinsically more difficult to analyze. Such vibrations certainly exist, and one would ideally desire to have as simple a description of their basic properties as is practical. For vibrations of complex systems, variational principles have been found to be highly useful for devising simpler analytical models. In this spirit, one finds frequent applications of Hamilton's equations, Lagrange equations derived from an approximated Lagrangian, Rayleigh's principle, and the Rayleigh–Ritz technique. However, the application of such concepts and techniques to the vibration of inclusions in extended elastic media is not immediately performed, since the mechanical system is unbounded. Viable analytical models with only a relatively small number of lumped parameters are desired, and one furthermore desires that these lumped parameters have a clear physical interpretation. The present paper describes one general methodology for achieving this. For simplicity, as well as to allow the use of relatively simple explicit examples, the analysis is limited to inclusions of a spherical shape.

2. ELASTODYNAMICS OF INHOMOGENEOUS MEDIA

The formulation here begins with that for an arbitrary orthogonal curvilinear coordinate system with the coordinates r, θ, and ϕ. The use of the symbols corresponding to spherical coordinates is deliberate, although the equations in this section hold for any choice of coordinate system, including the rectangular coordinate system.

Cauchy's equations of motion [17] for an isotropic solid can be derived from Hamilton's principle [18] (with the integration volume and the time integration kept deliberately vague, and with the absence of external forces and nonconservative forces in the region of integration)

$$\delta\{\iiiint \mathcal{L} h_r h_\theta h_\phi dr d\theta d\phi dt\} = 0, \tag{1}$$

where $\mathcal{L} = \mathcal{T} - \mathcal{V}$ is the Lagrangian per unit volume. Here, $\mathcal{T}$ is the kinetic energy per unit volume and $\mathcal{V}$ is the potential energy per unit volume; these are given by

$$\mathcal{T} = \frac{1}{2}\rho \frac{\partial \mathbf{u}}{\partial t}\frac{\partial \mathbf{u}}{\partial t}, \tag{2}$$

$$\mathcal{V} = \frac{1}{2}\lambda(\nabla \cdot \mathbf{u})^2 + \mu \sum_{\alpha,\beta} (\epsilon_{\alpha,\beta})^2. \tag{3}$$

Here, λ and μ are the Lamé coefficients and are possibly dependent on position. The quantities $\epsilon_{\alpha,\beta}$ are the strain-tensor components in a specified curvilinear orthogonal coordinate system. The length scales that appear in the volume integration in Eq. (1) are defined so that the square of the differential of distance is

$$(ds)^2 = h_r^2 (dr)^2 + h_\theta^2 (d\theta)^2 + h_\phi^2 (d\phi)^2. \tag{4}$$

The displacement vector $\mathbf{u}$ has components u_r, u_θ, and u_ϕ, so that

$$\mathbf{u} = u_r \mathbf{e}_r + u_\theta \mathbf{e}_\theta + u_\phi \mathbf{e}_\phi, \tag{5}$$

where $\mathbf{e}_r$, $\mathbf{e}_\theta$, and $\mathbf{e}_\phi$ are unit vectors pointing in the direction of the increase of the corresponding displacement coordinate.

In terms of the various quantities just introduced, a representative strain component is given by

$$2\epsilon_{r\theta} = \frac{1}{h_\theta}\frac{\partial u_r}{\partial \theta} - \frac{u_r}{h_\theta}\frac{\partial \mathbf{e}_r}{\partial \theta} \cdot \mathbf{e}_\theta + \frac{1}{h_r}\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{h_r}\frac{\partial \mathbf{e}_\theta}{\partial r} \cdot \mathbf{e}_e. \tag{6}$$

The other strain components are denoted analogously. The diagonal components can be regarded as incremental changes in length per unit length, and the off-diagonal components can be regarded as half the decrease in angle in radians between two lines originally aligned parallel to the local increasing directions of the coordinates.

In the application of Hamilton's principle under the circumstances just stated, one limits the admissible variations to be such that they vanish at all points on the boundaries of the fourfold integration region. This restriction is sufficient to derive partial differential equations (Lagrange–Euler equations) that hold within the interior of the region. There are three such equations; the equation corresponding to the r coordinate is

$$\begin{aligned} &\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{M}}{\partial\{\partial u_r/\partial t\}}\right) + \frac{\partial}{\partial r}\left(\frac{\partial \mathcal{M}}{\partial\{\partial u_r/\partial r\}}\right) \\ &+ \frac{\partial}{\partial \theta}\left(\frac{\partial \mathcal{M}}{\partial\{\partial u_r/\partial \theta\}}\right) + \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{M}}{\partial\{\partial u_r/\partial \phi\}}\right) - \frac{\partial \mathcal{M}}{\partial u_r} = 0, \end{aligned} \tag{7}$$

with the abbreviation $\mathcal{M} = \mathcal{L}h_r h_\theta h_\phi$. The product of the three length factors is recognized as the Jacobian. Analogous equations hold for the θ and ϕ components.

For the specific case of spherical coordinates, the length factors are $h_r = 1$, $h_\theta = r$, and $h_\phi = r\sin\theta$. The unit vectors are such that all are independent of r; the nonzero derivatives of these unit vectors with respect to the coordinates are

$$\frac{\partial \mathbf{e}_r}{\partial \theta} = \mathbf{e}_\theta; \quad \frac{\partial \mathbf{e}_\theta}{\partial \theta} = -\mathbf{e}_r, \tag{8}$$

$$\frac{\partial \mathbf{e}_r}{\partial \phi} = \mathbf{e}_\phi \sin\theta; \quad \frac{\partial \mathbf{e}_\theta}{\partial \phi} = \mathbf{e}_\phi \cos\theta, \tag{9}$$

$$\frac{\partial \mathbf{e}_\phi}{\partial \phi} = -\mathbf{e}_r \sin\theta - \mathbf{e}_\theta \cos\theta. \tag{10}$$

3. AXIALLY SYMMETRIC MOTIONS

The interest in the present paper involves those circumstances where, within the context of a system of spherical coordinates, the Lamé constants depend only on the radial coordinate r and where the overall disturbance is symmetric about the z axis, with $z = r\cos\theta$. The only two components of the particle displacement are u_r and u_θ, both of which depend in general on r, θ, and t but not on the axial coordinate ϕ. In such circumstances, the coupled partial differential equations derived in the manner described above can, after some algebra, be reduced to

$$\begin{aligned}\rho\frac{\partial^2 u_r}{\partial t^2} &= \frac{\partial \tau_{rr}}{\partial r} + 4\mu\frac{\partial}{\partial r}\left(\frac{u_r}{r}\right) \\ &+ \mu r^2 \Omega\frac{\partial}{\partial r}\left(\frac{u_\theta}{r^3}\right) + \frac{\mu}{r^2}\Omega\frac{\partial u_r}{\partial \theta},\end{aligned} \tag{11}$$

$$\begin{aligned}\rho\frac{\partial^2 u_\theta}{\partial t^2} &= \frac{1}{r^3}\frac{\partial (r^3\tau_{r\theta})}{\partial r} + \frac{2\mu}{r^2}u_\theta \\ &+ \frac{1}{r}\frac{\partial}{\partial \theta}\left(\frac{\lambda}{r}\frac{\partial}{\partial r}(r^2 u_r) + 2\mu u_r + (\lambda + 2\mu)\Omega u_\theta\right).\end{aligned} \tag{12}$$

Here, τ_{rr} and $\tau_{r\theta}$ are the stress components

$$\tau_{rr} = \lambda\nabla\cdot\mathbf{u} + 2\mu\frac{\partial u_r}{\partial r}, \tag{13}$$

$$\tau_{r\theta} = \mu\left(\frac{1}{r}\frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r} + \frac{\partial u_\theta}{\partial r}\right) \tag{14}$$

and Ω is the linear operator defined such that

$$\Omega F = \frac{1}{\sin\theta}\frac{\partial}{\partial \theta}(F\sin\theta) = \frac{\partial F}{\partial \theta} + (\cot\theta)F. \tag{15}$$

An important feature of these equations is that they are separable in the sense that, if $u_r(r, \theta, t)$ were expanded in terms of Legendre polynomials $P_n(\cos\theta)$ such that

$$u_r = \sum_{n=0}^{\infty} U_{r,n}(r,t)P_n(\cos\theta), \tag{16}$$

then the corresponding solution for the θ component has the form

$$u_\theta = \sum_{n=1}^{\infty} U_{\theta,n}(r,t)\frac{d}{d\theta}P_n(\cos\theta). \tag{17}$$

For any given n, the equations governing $U_{\theta,n}$ and $U_{r,n}$ are uncoupled from those for other values of n. This uncoupling occurs because the individual Legendre polynomials satisfy the differential equation

$$\Omega\frac{d}{d\theta}P_n(\cos\theta) + n(n+1)P_n(\cos\theta) = 0. \tag{18}$$

The residual equations that result when these expansions are inserted into the preceding equations are

$$\rho\frac{\partial^2 U_{r,n}}{\partial t^2} = \mathcal{Q}_{rr}\{U_{r,n}\} + \mathcal{Q}_{r\theta}\{U_{\theta,n}\}, \tag{19}$$

$$\rho\frac{\partial^2 U_{\theta,n}}{\partial t^2} = \mathcal{Q}_{\theta,n}\{U_{r,n}\} + \mathcal{Q}_{\theta\theta}\{U_{\theta,n}\}, \tag{20}$$

where

$$\begin{aligned}\mathcal{Q}_{rr}\{U_{r,n}\} &= \frac{\partial}{\partial r}\left[\frac{\lambda}{r^2}\frac{\partial}{\partial r}(r^2 U_{r,n}) + 2\mu\frac{\partial}{\partial r}U_{r,n}\right] \\ &+ 4\mu\frac{\partial}{\partial r}\left(\frac{U_{r,n}}{r}\right) - \frac{\mu}{r^2}n(n+1)U_{r,n},\end{aligned} \tag{21}$$

$$\begin{aligned}&\mathcal{Q}_{r\theta}\{U_{\theta,n}\} \\ &= -n(n+1)\left[\frac{\partial}{\partial r}\left(\frac{\lambda U_{\theta,n}}{r}\right) + \mu r^2\frac{\partial}{\partial r}\left(\frac{U_{\theta,n}}{r^3}\right)\right],\end{aligned} \tag{22}$$

$$\begin{aligned}&\mathcal{Q}_{\theta r}\{U_{r,n}\} \\ &= \frac{1}{r^3}\frac{\partial}{\partial r}(\mu r^2 U_{r,n}) + \frac{2\mu}{r^2}U_{r,n} + \frac{\lambda}{r^3}\frac{\partial}{\partial r}(r^2 U_{r,n}),\end{aligned} \tag{23}$$

$$\begin{aligned}\mathcal{Q}_{\theta\theta}\{U_{\theta,n}\} &= \frac{1}{r^3}\frac{\partial}{\partial r}\left[\mu r^4\frac{\partial}{\partial r}\left(\frac{U_{\theta,n}}{r}\right)\right] \\ &+ \frac{2\mu}{r^2}U_{\theta,n} - \frac{(\lambda+2\mu)}{r^2}n(n+1)U_{\theta,n}.\end{aligned} \tag{24}$$

These four linear operators $\mathcal{Q}_{rr}$, $\mathcal{Q}_{r\theta}$, etc., depend on the radial coordinate, the Lamé coefficients, and the index n, as is evident in the expressions above.

4. REDUCED VARIATIONAL PRINCIPLE

The residual equations just derived can also be obtained directly from Hamilton's principle, with the insertions

$$u_r \longrightarrow U_{r,n}(r,t)P_n(\cos\theta);$$
$$u_\theta \longrightarrow U_{\theta,n}\frac{d}{d\theta}P_n(\cos\theta). \tag{25}$$

For the application of Hamilton's principle, one specifies that the volume of interest is a sphere centered at the origin and with some fixed radius R. As before, admissible variations are restricted to be those that vanish at points on the integration boundary, including all points on the external spherical surface.

One then does the (trivial) integration over ϕ and the integration over θ. The kinetic energy, for example, becomes

$$\mathrm{KE} = \pi\int\rho\left[\left(\frac{\partial U_r}{\partial t}\right)^2 I_1 + \left(\frac{\partial U_\theta}{\partial t}\right)^2 I_2\right]r^2 dr, \tag{26}$$

where

$$I_1 = \int_0^\pi [P_n(\cos\theta)]^2 \sin\theta d\theta = \frac{2}{2n+1}, \tag{27}$$

$$I_2 = \int_0^\pi \left[\frac{d}{d\theta}P_n(\cos\theta)\right]^2 \sin\theta d\theta = \frac{2n(n+1)}{2n+1} \tag{28}$$

are definite integrals involving Legendre polynomials (here, for simplicity, we omit the subscripts n on $U_{r,n}$ and $U_{\theta,n}$, with the understanding that we are here dealing with a particular value of n).

Similarly, the contribution from the n terms to the total potential energy is

$$\mathrm{PE} = \pi\int[\lambda V_\lambda + 2\mu V_\mu]r^2 dr, \tag{29}$$

where

$$V_\lambda = \left(\frac{\partial U_r}{\partial r} + \frac{2}{r}U_r - \frac{n(n+1)}{r}U_\theta\right)^2 I_1, \tag{30}$$

$$V_\mu = \left[\left(\frac{\partial U_r}{\partial r}\right)^2 + \frac{2}{r^2}U_r^2\right]I_1 + \frac{2}{r^2}U_r U_\theta I_3 + \frac{1}{r^2}U_\theta^2 I_4 + \frac{1}{2r^2}\left(U_r - U_\theta + r\frac{\partial U_\theta}{\partial r}\right)^2 I_2. \tag{31}$$

Here, the integrals I_1 and I_2 are as given previously and

$$I_3 = \int_0^\pi P_n\frac{d^2P_n}{d\theta^2}\sin\theta d\theta + \int_0^\pi P_n\frac{dP_n}{d\theta}\cos\theta d\theta = -I_2, \tag{32}$$

$$I_4 = \int_0^\pi\left(\frac{d^2P_n}{d\theta^2}\right)^2\sin\theta d\theta + \int_0^\pi\left(\frac{dP_n}{d\theta}\right)^2\cot^2\theta\sin\theta d\theta = (n^2+n-1)I_2. \tag{33}$$

These integrals are, for the most part, derivable by integration by parts and by explicit use of ordinary differential equation (18) and of fundamental integrals given in basic texts [19].

One notes that the coupled partial differential equations (Eqs. (19) and (20)) are alternately derived from

$$\frac{\partial}{\partial t}\left(\frac{\partial\mathscr{P}}{\partial\{\partial U_r/\partial t\}}\right) + \frac{\partial}{\partial r}\left(\frac{\partial\mathscr{P}}{\partial\{\partial U_r/\partial r\}}\right) - \frac{\partial\mathscr{P}}{\partial U_r} = 0, \tag{34}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial\mathscr{P}}{\partial\{\partial U_\theta/\partial t\}}\right) + \frac{\partial}{\partial r}\left(\frac{\partial\mathscr{P}}{\partial\{\partial U_\theta/\partial r\}}\right) - \frac{\partial\mathscr{P}}{\partial U_\theta} = 0, \tag{35}$$

where

$$\mathscr{P} = \pi\rho\left[\left(\frac{\partial U_r}{\partial t}\right)^2 I_1 + \left(\frac{\partial U_\theta}{\partial t}\right)^2 I_2\right]r^2 - \pi[\lambda V_\lambda + 2\mu V_\mu]r^2 \tag{36}$$

is the apparent Lagrangian density per unit radial distance.

5. ASYMPTOTIC BEHAVIOR

The interest here is in dynamical behavior that is composed primarily of frequencies that are so low that the wavelengths of compressional and shear waves in the external elastic solid are much larger than any length scale associated with the inclusion. We let ω be such a frequency and let k_e and κ_e be the wave numbers for compressional and shear waves in the external medium. Then, $k_e a \ll 1$ and $\kappa_e a \ll 1$. One allows, however, for the possibility that the inclusion may have a characteristic time scale t_i that is such that, for some such frequencies, ωt_i is comparable to unity. This implies that one should not neglect the kinetic energy in the inner region, where r is on the order of a. One can nevertheless adopt the general philosophy of the method of matched asymptotic expansions [20] and assume that the inner solution for all such frequencies approaches a quasi-static solution of the residual equations at large values of r.

Within the outer region, the most general solution can be expressed as

$$U_{r,\,\mathrm{outer}} = \frac{\partial}{\partial r}\left[r^n\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n\left(\frac{F_A(t-[r/c_1])+F_B(t+[r/c_1])}{r}\right)\right] + n(n+1)r^{n-1}\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n\left(\frac{G_A(t-[r/c_2])+G_B(t+[r/c_2])}{r}\right), \tag{37}$$

$$U_{\theta,\,\mathrm{outer}} = r^{n-1}\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n\left(\frac{F_A(t-[r/c_1])+F_B(t+[r/c_1])}{r}\right) + \frac{1}{r}\frac{\partial}{\partial r}\left[r^{n+1}\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n\left(\frac{G_A(t-[r/c_2])+G_B(t+[r/c_2])}{r}\right)\right]. \tag{38}$$

Here, the arbitrary functions F_A and F_B correspond to outgoing and incoming longitudinal waves, while the functions G_A and G_B correspond to outgoing and incoming shear waves. All of these are functions of a single variable, while

$$c_1 = \left(\frac{\lambda_e + 2\mu_e}{\rho_e}\right)^{1/2}, \quad c_2 = \left(\frac{\mu_e}{\rho_e}\right)^{1/2} \tag{39}$$

are the wave speeds of longitudinal and shear waves in the external medium. The subscript "*e*" denotes values that correspond to the external medium. The derivation of the above relations is omitted for brevity, but can be deduced from the formulation in terms of potentials for a homogeneous medium given by Ying and Truell [4] and from Rayleigh's expression for the spherical Bessel functions.

To the lowest order in the parameters describing the size of the inclusion, the outer solution should be the same as if the inclusion were not present, so the corresponding expressions for U_r and U_θ should be finite at $r = 0$. To determine what restrictions this condition imposes on the functions that appear in the above expressions, one first notes that

$$r^n\left(-\frac{1}{r}\frac{d}{dr}\right)^n\left(\frac{r^s}{r}\right) = \frac{(1-s)(3-s)(5-s)\ldots(2n-1-s)}{r^{n+1-s}} \tag{40}$$

is zero if s is odd and if $1 < s < 2n - 1$. If s is odd and $s > 2n + 1$, then the above is zero (for $n \geq 1$) at $r = 0$. If, on the other hand, $s = 0$, then the expression is singular at $r = 0$. This requires that power series expansions of $F_A + F_B$ in r only contain odd powers of r, so that the overall expression is odd in r. This condition is met if one sets

$$F_A(t-[r/c_1]) + F_B(t+[r/c_1]) = F(t+[r/c_1]) - F(t-[r/c_1]), \tag{41}$$

where F is some given function, so that

$$F(t+[r/c_1]) - F(t-[r/c_1]) = 2(r/c_1)F'(t) + \frac{2}{3!}(r/c_1)^3F'''(t)\ldots + \frac{2}{(2n+1)!}(r/c_1)^{2n+1}F^{(2n+1)}(t) + \frac{2}{(2n+3)!}(r/c_1)^{2n+3}F^{(2n+3)}(t) + \ldots. \tag{42}$$

Here, $F^{(2n+1)}(t)$ is the $(2n + 1)$th derivative of $F(t)$.

In a similar manner, one can define a function $G(t)$ with regard to the shear wave terms.

In terms of the quantities just defined, one can write the two leading terms, valid for small r, of this outer solution as

$$U_{r,\,\mathrm{asymp}} = pnF_1(t)r^{n-1} + qF_2''(t)r^{n+1}, \tag{43}$$

$$U_{\theta,\,\mathrm{asymp}} = pF_1(t)r^{n-1} + qF_3''(t)r^{n+1}, \tag{44}$$

where

$$p = \frac{2}{(2n+1)!}[(-2n)(-2n+2)(-2n+4)\ldots(-2)], \tag{45}$$

$$q = \frac{2}{(2n+3)!}[(-2n-2)(-2n)(-2n+2)\ldots(-4)]. \tag{46}$$

The three time-dependent functions $F_1(t)$, $F_2(t)$, and $F_3(t)$ are related to the previously introduced quantities as

$$F_1(t) = \left(\frac{1}{c_1}\right)^{2n+1}F^{(2n+1)}(t) + (n+1)\left(\frac{1}{c_2}\right)^{2n+1}G^{(2n+1)}(t), \tag{47}$$

$$F_2(t) = (n+2)\left(\frac{1}{c_1}\right)^{2n+3}F^{(2n+1)}(t) + n(n+1)\left(\frac{1}{c_2}\right)^{2n+3}G^{(2n+1)}(t), \tag{48}$$

$$F_3(t) = \left(\frac{1}{c_1}\right)^{2n+3} F^{(2n+1)}(t) + (n+3)\left(\frac{1}{c_2}\right)^{2n+3} G^{(2n+1)}(t). \quad (49)$$

Only two of these functions are independent. The derivable relationship is

$$(4n+6)F_1 = [(n+3)c_1^2 - (n+1)c_2^2]F_2 - (n+1)[nc_1^2 - (n+2)c_2^2]F_3. \quad (50)$$

In the terminology of the method of matched asymptotic expansions [20], the asymptotic expressions $U_{r,\,\mathrm{asymp}}$ and $U_{\theta,\,\mathrm{asymp}}$ are the leading terms in the inner expansion of the lowest order term of the outer solution.

(The remainder of the manuscript is being polished.)

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity*, 3rd ed. (Nauka, Moscow, 1965; Pergamon, Oxford, 1970) (The result of interest is attributed to work in 1949 by M. A. Isakovich).
2. E. Meyer, K. Brendel, and K. Tamm, J. Acoust. Soc. Am. **30**, 1116 (1958).
3. D. V. Sivukhin, Sov. Phys. Acoust. **1**, 82 (1955).
4. C. F. Ying and R. Truell, J. Appl. Phys. **27**, 1086 (1956).
5. K. F. Herzfeld, Philos. Mag. **9**, 741 (1930).
6. L. Knopoff, Geophysics **24**, 30 (1959).
7. N. G. Einspruch and R. Truell, J. Acoust. Soc. Am. **32**, 214 (1960).
8. G. Johnson and R. Truell, J. Appl. Phys. **36**, 3466 (1965).
9. C. C. Mow, J. Appl. Mech. **32**, 637 (1965).
10. F. R. Norwood and J. Miklowitz, J. Appl. Mech. **34**, 735 (1967).
11. G. C. Gaunaurd and H. Überall, J. Acoust. Soc. Am. **63**, 1699 (1978).
12. V. N. Alekseev and S. A. Rybak, Acoust. Phys. **45**, 535 (1999).
13. J. S. Allen and R. A. Roy, J. Acoust. Soc. Am. **107**, 3167 (2000).
14. D. B. Khismatullin and A. Nadim, Phys. Fluids **14**, 3534 (2002).
15. Yu. M. Zaslavskiĭ, Acoust. Phys. **50**, 46 (2004).
16. S. Y. Emelianov, M. F. Hamilton, Y. A. Ilinskii, and E. A. Zabolotskaya, J. Acoust. Soc. Am. **115**, 581 (2004).
17. I. S. Sokolnikoff, *Mathematical Theory of Elasticity* (McGraw-Hill, New York, 1956), pp. 177–184.
18. A. D. Pierce, in *Physical Acoustics: Principles and Methods*, Ed. by W. P. Mason (Academic, San Diego, 1984; Mir, Moscow, 1984), Vol. 22, pp. 195–371.
19. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. (Cambridge Univ. Press, Cambridge, 1927), p. 309.
20. M. van Dyke, *Perturbation Methods in Fluid Mechanics* (Academic, New York, 1964; Mir, Moscow, 1967), pp. 77–97.
21. J. N. Goodier, J. Appl. Mech., Paper **APM-55-7**, 39 (1933).
22. V. C. Anderson, J. Acoust. Soc. Am. **22**, 426 (1950).

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 52–62. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 68–80.

Spectral–Morphological Analysis of Acoustical Images of Biological Tissues and Composite Structures: I. Statistical Approach

V. A. Burov, E. L. Kim, and O. D. Rumyantseva

Faculty of Physics, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia

e-mail: burov@phys.msu.ru

Received June 16, 2004

Abstract—The problem of classifying images of different biological tissues and composite structures is solved using the spectral and morphological analysis based on the Bayesian method for statistical hypothesis verification. The basis functions are constructed from a learning set. The spectral approach and its particular realizations in the form of Bartlett's and Pisarenko's methods adapted to the problem are considered. An extension of the spectral approach to the more general spectral–morphological classification is proposed. The latter takes into account the spatial-spectrum features of the structure types to be classified, as well as their morphological features, which manifest themselves in a correlation between the expansion coefficients. The characteristic properties of the spectral and spectral–morphological approaches are discussed using numerical classification examples. The method is generalized to the classification of multiparameter images of structures, which may be represented, for example, by the distributions of the sound velocity, density, absorption, and values of the nonlinear parameter. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

In this paper, we solve the problem of the statistical classification of images of biological tissues by spectral and morphological methods. For example, the spatial distributions obtained for the quantitative characteristics (including multiparameter ones) of tissues from acoustoscopic experiments need a further secondary classification in order to make an overall diagnostic decision. A review of the mathematical foundations of digital image processing developed to date is given in [1]. At the same time, the classification methods for tomographic and microscopic images of composite structures and biological tissues can be categorized from a physical viewpoint into three classes, which, however, have no clear-cut boundaries between them: locally parametric methods, nonlocal spectral statistical methods, and morphological methods. There are also publications devoted to image-quality improvement with the aim of enhancing visual perception [2, 3]. The first class includes locally parametric methods and suggests that a combination of quantitative characteristics (parameters such as, for example, velocity of sound and attenuation) be analyzed at one point; i.e., the characterization is purely local [4]. In [5, 6], such parameters as attenuation and backscattering coefficient are estimated in parallel. However, the backscattering coefficient cannot be attributed to a particular point and implies a certain structured scattering volume whose size is comparable with the wavelength. Therefore, [5, 6] in essence represent a transition from the first class of local methods to the second class. Nonlocal spectral statistical methods, which constitute the second class (the largest in terms of the number of publications), assume classification according to the statistical characteristics of an isolated small image area or a tissue cut. This class includes the spectral methods that process the acoustic echo signal carrying the information on the object [7–10] and nonlocal classification methods that are based on the brightness distribution function [11]. In addition, one of the fundamental problems of image analysis is the development of mathematical description methods that convey the image content and meaning. The description must only convey the image features that are essential from the viewpoint of the problem being solved and must be independent of inessential features. In this regard, a great number of publications are devoted to developing the statistical properties of images [12, 13], so that these properties can also be used in the problems of classification based on neural networks [14, 15].

The third class is represented by morphological methods, which analyze the internal structure of the image. For the morphological analysis, conditions of recording the object's image and the parameters of the recording equipment are insignificant characteristics [16]. For images of engineering structures, these methods provide good results, because the clear-cut shape of the image being sought is known a priori. In these problems, the classification often consists of the minimization of a discrepancy with respect to linear dimensions or orientation of an object of a known shape [16–18]. However, a disadvantage of the purely morphological

analysis in the classification of biological tissues is that, as a rule, the exact shape of the tissue structure element being classified is initially unknown (for example, there exist a great number of cell formations even for the one type of a mammary cancer).

This paper develops classification methods for images of structures and biological tissues that can be used to separate regions with particular statistical, parametric, and morphological properties. These methods can be used to process the images independently of the technique used to produce them, as well as to directly classify the structures on the basis of experimental acoustic scattering data. The approaches are applicable to one-parameter and multiparameter classification problems with the parametric, spectral, and morphological aspects of the problem taken into account sequentially or simultaneously. In this respect, they hold an intermediate and, at the same time, unifying position between the classes mentioned above. We also generalize the analysis to multicomponent images, in which the components are the tomographic data, for example, the velocity of sound, the attenuation, and the nonlinear parameter. Thus, the specificity of the problem under consideration, as applied to acoustic images produced by tomographs, acoustic microscopes, etc., consists in the fundamental possibility of performing a multiparameter classification of tissues in combination with the observed morphological features of their structure.

2. CLASSIFICATION IN TERMS OF THE LIKELIHOOD RATIO

The general approach proposed here can, to a certain extent, be conventionally called the optimal one, because it relies on the Bayesian method of statistical hypothesis verification, i.e., on the likelihood ratio [1]. Particular criteria that follow from this general approach refer to spatial multidimensional spectral–correlation analysis and, in this sense, are close to the algorithms developed earlier [7, 10]. The difference is that our approach employs a sequence of linear and nonlinear transformations, which sharpen the response of the algorithms to a spatial region with a given structure.

The physical essence of the approach is as follows. The brightness image to be classified is described by a two-dimensional nonnegative function $\Gamma_0(\mathbf{r})$. In this image, it is necessary to find regions with structural features of a given type. Each type is greatly determined by the shape and size of its characteristic elements and also by their relative arrangement. Let two types of structure be defined a priori, which we label with indices I and II. Then, a criterion of whether or not each current fragment $X_\mathbf{r}$ ($\mathbf{r}$ is the center of the fragment) of the image $\Gamma_0(\mathbf{r})$ belongs to one of the given types is constructed at each particular point $\mathbf{r}$ based on the current likelihood ratio function $L(X_\mathbf{r}) \equiv L(\mathbf{r}) = P(X_\mathbf{r}|\mathrm{I})/P(X_\mathbf{r}|\mathrm{II})$. Here, $P(X_\mathbf{r}|\mathrm{I})$ and $P(X_\mathbf{r}|\mathrm{II})$ are the conditional probability densities of the event that the fragment X_r belongs to the structure of the first type (this assumption is the hypothesis) and of the second type (this is the alternative), respectively. A close approach is applied in [19] to a similar problem of classification of random noise signals in terms of a specified shape of their spectrum and of simultaneous estimation of the spectrum parameters. However, the problem of image classification has several specific features. First, the images are two-dimensional. Second, the space in which the images are classified in our approach is the eigenvector space of the autocovariance matrix, which is specified a priori or estimated for each structure type. From the viewpoint of the initial information, this is equivalent to classification based on differences in the local spatial power spectrum, with the spectrum in this case being known completely rather than being accurate to unknown parameters, as in [19].

We assume that a learning set is given for each type of structure. Each sample of this set is described by the two-dimensional function $u(\mathbf{r})$. On the one hand, the size of the sample is much smaller than that of the whole image. On the other hand, the sample must contain specific features of the structure type being classified. For each of the given structure types, a basis that describes its statistical properties is constructed. We use the Karhunen–Loéve basis [1]. It is constructed with the help of the autocovariance matrix $A(\mathbf{r}, \mathbf{r}')$ of an ensemble of functions $u(\mathbf{r})$: $A(\mathbf{r}, \mathbf{r}') \equiv \overline{u_{\mathrm{dif}}(\mathbf{r})u^*_{\mathrm{dif}}(\mathbf{r}')}$, where $u_{\mathrm{dif}}(\mathbf{r}) \equiv u(\mathbf{r}) - \bar{u}$ and $\bar{u} \equiv \overline{u(\mathbf{r})}$. Here and below, the overbar symbol means averaging over the ensemble of realizations of the given type; the asterisk symbol means complex conjugation (for the analysis to be general, complex designations are used). Subtracting the average brightnesses $\bar{\Gamma}_0$ and $\bar{u}$ from the whole image $\Gamma_0(\mathbf{r})$ being classified and from each sample $u(\mathbf{r})$ (respectively) removes the effect of brightness on the classification process. The values $\bar{\Gamma}_0$ and $\bar{u}$ can be calculated by averaging over a series of images and series of samples or by averaging $\Gamma_0(\mathbf{r})$ and $u(\mathbf{r})$ over coordinates (which we used in numerical simulations under the assumption of the spatial ergodicity). As a result, we obtain an alternating-sign image $\Gamma_{\mathrm{dif}}(\mathbf{r}) \equiv \Gamma_0(\mathbf{r}) - \bar{\Gamma}_0$ with a zero mean.

The exact form of the autocovariance matrix $A(\mathbf{r}, \mathbf{r}')$ for the complete ensemble with given correlation properties to which the learning set belongs is unknown. Therefore, we use an estimate of this matrix obtained under the assumption that the statistical properties of the structure being classified are spatially uniform. According to this assumption, the autocovariance operator is constructed as a Hermitian block Toeplitz matrix: $A(\mathbf{r}, \mathbf{r}') = A(\mathbf{r}' - \mathbf{r}) = A^*(\mathbf{r} - \mathbf{r}')$. In this manner, functions of the autocovariance cut $K_{uu}(\boldsymbol{\rho}) \equiv \frac{1}{V_u}\int u_{\mathrm{dif}}(\mathbf{r})u^*_{\mathrm{dif}}(\mathbf{r} + \boldsymbol{\rho})d\mathbf{r}$ are constructed one by one for each sample $u(\mathbf{r})$ of

the learning set. The normalizing factor V_u is equal to the area occupied by the sample. The functions $K_{uu}(\boldsymbol{\rho})$ are averaged over the samples: $K(\boldsymbol{\rho}) \equiv \overline{K_{uu}}(\boldsymbol{\rho})$. Finally, the matrix $A(\mathbf{r}, \mathbf{r}')$ is constructed as described by the relationship

$$A(\mathbf{r}, \mathbf{r}') \approx K(\boldsymbol{\rho} = \mathbf{r}' - \mathbf{r}) \tag{1}$$

under the assumption of spatial homogeneity and ergodicity. The following analysis relies on the hypothesis that this matrix is associated with a certain complete ensemble, which below will be called the hypothetical ensemble. Such an ensemble consists of a set of all spatially homogeneous (in the statistical sense) images whose correlation properties are described by the matrix $A(\mathbf{r}, \mathbf{r}')$ obtained above. The set that was used to estimate $A(\mathbf{r}, \mathbf{r}')$ is a subset of this complete hypothetical ensemble.

The Karhunen–Loéve basis consists of eigenvectors $\varphi(\mathbf{r})$ of the matrix $A(\mathbf{r}, \mathbf{r}')$, which correspond to eigenvalues λ and, by definition, satisfy the relationship

$$\int A(\mathbf{r}, \mathbf{r}')\varphi(\mathbf{r}')d\mathbf{r}' = \lambda\varphi(\mathbf{r}). \tag{2}$$

All λ are real, because the matrix $A(\mathbf{r}, \mathbf{r}')$ is Hermitian, and are nonnegative because of its autocorrelation origin. When we change from the continuous functions to their discrete analogs, the sample functions $u(\mathbf{r})$ are replaced with their values at $N \times N$ discrete points. Then, the dimension of the block Toeplitz matrix $A(\mathbf{r}, \mathbf{r}')$ is $N^2 \times N^2$.

In subsequent calculations, it is convenient to use the Dirac notation. Discrete values of the samples $u_{\text{dif}}(\mathbf{r})$ without the average value are represented as a column vector consisting of N^2 components: $u_{\text{dif}} \equiv |T^{\text{I}}\rangle$ and $u_{\text{dif}} \equiv |T^{\text{II}}\rangle$ for structure types I and II, respectively. In this notation, relationship (2) takes the following form (the discrete analog of the spatial element $d\mathbf{r}$ is assumed to be equal to unity):

$$A^{\text{I}}|\varphi_i\rangle = \lambda_i|\varphi_i\rangle, \quad A^{\text{II}}|\psi_j\rangle = \mu_j|\psi_j\rangle, \quad i, j = 1, \ldots, N^2. \tag{3}$$

Here, $\{\varphi_i\}$ and $\{\lambda_i\}$ are the sets of orthonormal eigenvectors and eigenvalues (numbered by index i) for structure type I (i.e., $A = A^{\text{I}}$); similarly, $\{\psi_j\}$ and $\{\mu_j\}$ are those for structure type II (i.e., $A = A^{\text{II}}$). The initial format of the samples $u(\mathbf{r})$ in the sequences of both types are assumed to be the same and, therefore, the number of components in the eigenvectors φ_i and ψ_j is the same.

The basis $\{\varphi_i\}$ is the proper basis for the structure of the first type $|T^{\text{I}}\rangle$ and a "foreign" alternative basis for the structure of the second type $|T^{\text{II}}\rangle$. For the basis $\{\psi_j\}$, the situation is reversed. These bases are used below to construct the likelihood ratio $L(\mathbf{r})$. Then, coefficients of expansion of the structure of a given type in its proper basis and in the alternative basis are

$$a_i^{\text{I}} \equiv \langle\varphi_i|T^{\text{I}}\rangle, \quad b_j^{\text{I}} \equiv \langle\psi_j|T^{\text{I}}\rangle;$$
$$a_j^{\text{II}} \equiv \langle\psi_j|T^{\text{II}}\rangle, \quad b_i^{\text{II}} \equiv \langle\varphi_i|T^{\text{II}}\rangle. \tag{4}$$

In a similar way, the coefficients of expansion of the vector $|X_{\mathbf{r}}\rangle$, which is a current fragment of the whole image $\Gamma_{\text{dif}}(\mathbf{r})$ being classified, are introduced:

$$c_i^{\text{I}}(\mathbf{r}) \equiv \langle\varphi_i|X_{\mathbf{r}}\rangle \equiv \int X_{\mathbf{r}}(\mathbf{r}')\varphi_i^*(\mathbf{r}')d\mathbf{r}',$$
$$c_j^{\text{II}}(\mathbf{r}) \equiv \langle\psi_j|X_{\mathbf{r}}\rangle \equiv \int X_{\mathbf{r}}(\mathbf{r}')\psi_j^*(\mathbf{r}')d\mathbf{r}'. \tag{5}$$

For each type of structure, we construct correlation matrices of expansion coefficients in the proper basis (matrix K_{aa}), in the alternative basis (matrix K_{bb}), and in both bases simultaneously (matrices K_{ab} and K_{ba} of cross coefficients). The elements of these matrices (described by indices $m, n = 1, \ldots, N^2$) for structure $|T^{\text{I}}\rangle$ are as follows:

$$(K_{aa}^{\text{I}})_{mn} \equiv \overline{a_m^{\text{I}}(a_n^{\text{I}})^*} = \overline{\langle\varphi_m|T^{\text{I}}\rangle\langle T^{\text{I}}|\varphi_n\rangle};$$
$$(K_{bb}^{\text{I}})_{mn} \equiv \overline{b_m^{\text{I}}(b_n^{\text{I}})^*} = \overline{\langle\psi_m|T^{\text{I}}\rangle\langle T^{\text{I}}|\psi_n\rangle};$$
$$(K_{ab}^{\text{I}})_{mn} \equiv \overline{a_m^{\text{I}}(b_n^{\text{I}})^*} = \overline{\langle\varphi_m|T^{\text{I}}\rangle\langle T^{\text{I}}|\psi_n\rangle}; \tag{6}$$
$$(K_{ba}^{\text{I}})_{mn} \equiv \overline{b_m^{\text{I}}(a_n^{\text{I}})^*} = (K_{ab}^{\text{I}})_{nm}^*.$$

The averaging is performed over an ensemble of structures of the corresponding type. The purely spectral approach classifies structures only in terms of the spatial-spectrum features, and the above ensemble is the complete hypothetical ensemble corresponding to the estimated matrix A. The spectral–morphological approach, which is considered in Section 4 below, additionally takes into account the morphological features, which narrows the ensemble. In both cases, when we consider all N^2 eigenvectors and eigenvalues of each type, an explicit relationship between the matrices occurs: $K_{ab}^{\text{I}} = K_{aa}^{\text{I}}K_{\varphi\psi}$; $K_{ba}^{\text{I}} = (K_{ab}^{\text{I}})^+ = K_{\psi\varphi}K_{aa}^{\text{I}}$; $K_{bb}^{\text{I}} = K_{\psi\varphi}K_{ab}^{\text{I}}$, where the matrix $K_{\varphi\psi}$ with elements $(K_{\varphi\psi})_{i,j} \equiv \langle\varphi_i|\psi_j\rangle$ is the transition matrix between the bases $\{\varphi_i\}$ and $\{\psi_j\}$; $K_{\psi\varphi} = K_{\varphi\psi}^+$; and the plus symbol means Hermitian conjugation. For the structure $|T^{\text{II}}\rangle$, the corresponding elements $(K^{\text{II}})_{mn}$ are obtained from $(K^{\text{I}})_{mn}$ by replacing T^{I} with T^{II}, φ_m with ψ_m, φ_n with ψ_n, and matrices $K_{\varphi\psi}$ with $K_{\psi\varphi}$.

The resultant combined matrices $K^{\mathrm{I}}_{\mathrm{comb}}$ and $K^{\mathrm{II}}_{\mathrm{comb}}$ (their dimension is $2N^2 \times 2N^2$) for the structure of the corresponding type consist of four blocks:

$$K^{\mathrm{I}}_{\mathrm{comb}} \equiv \begin{pmatrix} K^{\mathrm{I}}_{aa};\ K^{\mathrm{I}}_{ab} \\ K^{\mathrm{I}}_{ba};\ K^{\mathrm{I}}_{bb} \end{pmatrix}, \quad K^{\mathrm{II}}_{\mathrm{comb}} \equiv \begin{pmatrix} K^{\mathrm{II}}_{aa};\ K^{\mathrm{II}}_{ab} \\ K^{\mathrm{II}}_{ba};\ K^{\mathrm{II}}_{bb} \end{pmatrix}. \quad (7)$$

These matrices, which are quadratic forms of the expansion coefficients, account for the individual statistical properties of structures of both types and for their "cross" properties. The current image fragment $|X_{\mathbf{r}}\rangle$ in the classification process is characterized by the column vectors $|(X_{\mathbf{r}})^{\mathrm{I}}_{\mathrm{II}}\rangle$ and $|(X_{\mathbf{r}})^{\mathrm{II}}_{\mathrm{I}}\rangle$:

$$|(X_{\mathbf{r}})^{\mathrm{I}}_{\mathrm{II}}\rangle = \begin{pmatrix} \{c^{\mathrm{I}}_i\}_{i=1,\ldots,N^2} \\ \{c^{\mathrm{II}}_j\}_{j=1,\ldots,N^2} \end{pmatrix},$$
$$|(X_{\mathbf{r}})^{\mathrm{II}}_{\mathrm{I}}\rangle = \begin{pmatrix} \{c^{\mathrm{II}}_j\}_{j=1,\ldots,N^2} \\ \{c^{\mathrm{I}}_i\}_{i=1,\ldots,N^2} \end{pmatrix}. \quad (8)$$

It was found numerically that real brightness histograms of the classified structures can be approximately described by a Gaussian distribution. Then, under the assumption that the random vectors $|(X_{\mathbf{r}})^{\mathrm{I}}_{\mathrm{II}}\rangle$ and $|(X_{\mathbf{r}})^{\mathrm{II}}_{\mathrm{I}}\rangle$ obey the multidimensional zero-mean Gaussian distribution, the classification likelihood ratio takes the form

$$L(\mathbf{r}) = \frac{P(X_{\mathbf{r}}|\mathrm{I})}{P(X_{\mathbf{r}}|\mathrm{II})}$$
$$= \frac{\det^{1/2}(K^{\mathrm{II}}_{\mathrm{comb}})\exp\left(-\frac{1}{2}\langle(X_{\mathbf{r}})^{\mathrm{I}}_{\mathrm{II}}|(K^{\mathrm{I}}_{\mathrm{comb}})^{-1}|(X_{\mathbf{r}})^{\mathrm{I}}_{\mathrm{II}}\rangle\right)}{\det^{1/2}(K^{\mathrm{I}}_{\mathrm{comb}})\exp\left(-\frac{1}{2}\langle(X_{\mathbf{r}})^{\mathrm{II}}_{\mathrm{I}}|(K^{\mathrm{II}}_{\mathrm{comb}})^{-1}|(X_{\mathbf{r}})^{\mathrm{II}}_{\mathrm{I}}\rangle\right)}.$$

The classification criterion is the result of a comparison of the function $L(\mathbf{r})$ with a given threshold value $(L_0)^{\mathrm{I}}_{\mathrm{II}}$ determined by the costs of the errors and by the prior-occurrence probabilities for structures of both types. It is convenient to make a decision by taking a logarithm $\ln L(\mathbf{r}) \equiv \Lambda(\mathbf{r})$ and comparing the following quantity with zero:

$$\Lambda'(\mathbf{r}) \equiv \Lambda(\mathbf{r}) - \ln(L_0)^{\mathrm{I}}_{\mathrm{II}}$$
$$= \frac{1}{2}[\langle(X_{\mathbf{r}})^{\mathrm{II}}_{\mathrm{I}}|(K^{\mathrm{II}}_{\mathrm{comb}})^{-1}|(X_{\mathbf{r}})^{\mathrm{II}}_{\mathrm{I}}\rangle$$
$$-\langle(X_{\mathbf{r}})^{\mathrm{I}}_{\mathrm{II}}|(K^{\mathrm{I}}_{\mathrm{comb}})^{-1}|(X_{\mathbf{r}})^{\mathrm{I}}_{\mathrm{II}}\rangle$$
$$+\ln(\det K^{\mathrm{II}}_{\mathrm{comb}}) - \ln(\det K^{\mathrm{I}}_{\mathrm{comb}})] - \ln(L_0)^{\mathrm{I}}_{\mathrm{II}}.$$

If the decision is made based on the two-level (binary) scheme, we assume that, if $\Lambda'(\mathbf{r}) > 0$, the fragment $X_{\mathbf{r}}$ is a type I structure or, if $\Lambda'(\mathbf{r}) < 0$, it is a type II structure.

General expressions (6) can be used and rendered concrete for the spatial–spectral approach, in which the averaging is performed over the complete hypothetical ensemble. By definition, this ensemble includes all structures whose correlation properties are determined by the given autocovariance matrix of the type under consideration, i.e., A^{I} or A^{II}. Therefore, here,

$$\overline{|T^{\mathrm{I}}\rangle\langle T^{\mathrm{I}}|} = A^{\mathrm{I}}, \quad \overline{|T^{\mathrm{II}}\rangle\langle T^{\mathrm{II}}|} = A^{\mathrm{II}}, \quad (9)$$

and Eq. (3) takes the form

$$\overline{|T^{\mathrm{I}}\rangle\langle T^{\mathrm{I}}|}\varphi_i\rangle = \lambda_i|\varphi_i\rangle, \quad \overline{|T^{\mathrm{II}}\rangle\langle T^{\mathrm{II}}|}\psi_j\rangle = \mu_j|\psi_j\rangle, \quad i, j = 1, \ldots, N^2. \quad (10)$$

Then, by virtue of property (10) and because the eigenvectors are orthonormal, the correlation matrices of expansion coefficients of the given type of structure in the proper basis, K^{I}_{aa} and K^{II}_{aa} of Eqs. (6), become diagonal:

$$(K^{\mathrm{I}}_{aa})_{mn} = \langle\varphi_m|\overline{|T^{\mathrm{I}}\rangle\langle T^{\mathrm{I}}|}\varphi_n\rangle = \lambda_n\delta_{mn};$$
$$(K^{\mathrm{II}}_{aa})_{mn} = \langle\psi_m|\overline{|T^{\mathrm{II}}\rangle\langle T^{\mathrm{II}}|}\psi_n\rangle = \mu_n\delta_{mn}; \quad (11)$$
$$\delta_{mn} \equiv \begin{cases} 1, & m = n \\ 0, & m \neq n. \end{cases}$$

Expressions for the cross coefficients are reduced to

$$(K^{\mathrm{I}}_{ab})_{mn} = \lambda_m\langle\varphi_m|\psi_n\rangle; \quad (K^{\mathrm{II}}_{ab})_{mn} = \mu_m\langle\psi_m|\varphi_n\rangle. \quad (12)$$

Finally, the relationship $\sum_{i=1}^{N^2}|\varphi_i\rangle\langle\varphi_i| = \hat{E}$ (and a similar expression for ψ_j), where $\hat{E}$ is the unit matrix, and property (10) yield

$$(K^{\mathrm{I}}_{bb})_{mn} = \sum_{i=1}^{N^2}\lambda_i\langle\psi_m|\varphi_i\rangle\langle\varphi_i|\psi_n\rangle;$$
$$(K^{\mathrm{II}}_{bb})_{mn} = \sum_{j=1}^{N^2}\mu_j\langle\varphi_m|\psi_j\rangle\langle\psi_j|\varphi_n\rangle. \quad (13)$$

It should be noted that the combination of the bases $\{\varphi_i\}_{i=1,\ldots,N^2}$ and $\{\psi_j\}_{j=1,\ldots,N^2}$ creates a redundant basis. In this regard, it is reasonable to discuss the classification process that uses a redundant basis and a nonredundant one. The technique proposed above for constructing the likelihood ratio function $L(\mathbf{r})$ with the help of the combined matrices $K^{\mathrm{I}}_{\mathrm{comb}}$ and $K^{\mathrm{II}}_{\mathrm{comb}}$ relies on the two bases simultaneously in the form of the combined

basis $\{\varphi_i\}_{i=1,\ldots,N^2} \cup \{\psi_j\}_{j=1,\ldots,N^2}$ used to construct each of the probability densities $P(X_\mathbf{r}|\mathrm{I})$ and $P(X_\mathbf{r}|\mathrm{II})$. In most situations, this basis is redundant, because, if we consider all N^2 eigenvectors of each type with all nonzero eigenvalues, each of the bases $\{\varphi_i\}$ and $\{\psi_j\}$ ($i, j = 1, \ldots, N^2$) is complete for describing any image fragment (whose spatial format is the same as that of the eigenvector), including structures of both types. If we take into account only $i = 1, \ldots, N_\mathrm{I}'$ and $j = 1, \ldots, N_\mathrm{II}'$ eigenvectors of type I and II that correspond to relatively large λ_i and μ_j (λ_i and μ_j are numbered in decreasing order of magnitude; N_I', $N_\mathrm{II}' < N$), the combined basis $\{\varphi_i\}_{i=1,\ldots,N_\mathrm{I}'} \cup \{\psi_j\}_{j=1,\ldots,N_\mathrm{II}'}$ is sufficient to describe structures of both types with a given relative accuracy δ, which can be estimated using the generalized Parseval's theorem: $\overline{\langle T^\mathrm{I}|T^\mathrm{I}\rangle} \equiv \sum_{i=1}^{N^2} \overline{\langle T^\mathrm{I}|\varphi_i\rangle\langle\varphi_i|T^\mathrm{I}\rangle} = \sum_i^{N^2} \overline{|a_i^\mathrm{I}|^2}$. Therefore, the relative accuracy δ_I, within which a type I structure can be described in terms of the truncated basis $\{\varphi_i\}_{i=1,\ldots,N_\mathrm{I}'}$ taking into account no less than $\cong|a_{i=N_\mathrm{I}'}^\mathrm{I}|$ terms, and, similarly, the accuracy δ_II, within which a type II structure can be described in terms of the truncated basis $\{\psi_j\}_{j=1,\ldots,N_\mathrm{II}'}$ taking into account no less than $\cong|a_{j=N_\mathrm{II}'}^\mathrm{II}|$ terms, are as follows: $\delta_\mathrm{I} \cong \sqrt{\sum_{i=N_\mathrm{I}'+1}^{N^2} \overline{|a_i^\mathrm{I}|^2} / \sum_{i=1}^{N^2} \overline{|a_i^\mathrm{I}|^2}}$ and $\delta_\mathrm{II} \cong \sqrt{\sum_{j=N_\mathrm{II}'+1}^{N^2} \overline{|a_j^\mathrm{II}|^2} / \sum_{j=1}^{N^2} \overline{|a_j^\mathrm{II}|^2}}$. In the purely spectral approach, $\overline{|a_i^\mathrm{I}|^2} \equiv (K_{aa}^\mathrm{I})_{ii} = \lambda_i$ and $\overline{|a_j^\mathrm{II}|^2} \equiv (K_{aa}^\mathrm{II})_{jj} = \mu_j$; then, $\delta_\mathrm{I} \cong \sqrt{\sum_{i=N_\mathrm{I}'+1}^{N^2} \lambda_i / \sum_{i=1}^{N^2} \lambda_i}$ and $\delta_\mathrm{II} \cong \sqrt{\sum_{j=N_\mathrm{II}'+1}^{N^2} \mu_j / \sum_{j=1}^{N^2} \mu_j}$, or coarser estimates are $\delta_\mathrm{I} \cong \sqrt{\lambda_\mathrm{min}/\lambda_\mathrm{max}}$ and $\delta_\mathrm{II} \cong \sqrt{\mu_\mathrm{min}/\mu_\mathrm{max}}$, where $\lambda_\mathrm{max} \equiv \lambda_{i=1}$ and $\mu_\mathrm{max} \equiv \mu_{j=1}$ are the maximum eigenvalues and λ_min and μ_min are the minimum eigenvalues taken into account and corresponding to the numbers $i = N_\mathrm{I}'$ and $j = N_\mathrm{II}'$. It is natural to take N_I' and N_II' such that the accuracies δ within which structures of both types are described are equal: $\delta = \delta_\mathrm{I} \approx \delta_\mathrm{II}$. The results of structure classification for model and real images, which we intend to discuss in the next paper, have shown that it is the redundant combined basis that is convenient to use.

Of course, a different technique can be used to define the ratio $L(\mathbf{r})$, for which $P(X_\mathbf{r}|\mathrm{I})$ is formed on the nonredundant basis $\{\varphi_i\}_{i=1,\ldots,N^2}$, while $P(X_\mathbf{r}|\mathrm{II})$, on $\{\psi_j\}_{j=1,\ldots,N^2}$. In this description, instead of Eqs. (8), we obtain $|(X_\mathbf{r})_\mathrm{II}^\mathrm{I}\rangle = \{c_i^\mathrm{I}\}_{i=1,\ldots,N^2}$ and $|(X_\mathbf{r})_\mathrm{I}^\mathrm{II}\rangle = \{c_j^\mathrm{II}\}_{j=1,\ldots,N^2}$, while the combined matrices $K_\mathrm{comb}^\mathrm{I}$ and $K_\mathrm{comb}^\mathrm{II}$ are replaced with matrices K_{aa}^I and K_{aa}^II, respectively. If only $i = 1, \ldots, N_\mathrm{I}'$ and $j = 1, \ldots, N_\mathrm{II}'$ eigenvectors are used, the following consideration must be taken into account. The truncated basis $\{\varphi_i\}_{i=1,\ldots,N_\mathrm{I}'}$ or $\{\psi_j\}_{j=1,\ldots,N_\mathrm{II}'}$ is complete for describing a structure of the corresponding (proper) type within the accuracy δ, but this basis may prove to be incomplete for describing a structure of the alternative type within the same accuracy. Therefore, in the method of constructing $L(\mathbf{r})$ considered here, the truncated bases must be augmented. To this end, only those vectors $\{\varphi_{i=i'}\}$ and $\{\psi_{j=j'}\}$ that describe the structure of the alternative type within an accuracy no higher than the accuracy within which the vector $\varphi_{i=N_\mathrm{I}'}$ describes the structure type I and the vector $\psi_{j=N_\mathrm{II}'}$ describe the structure type II are selected from the eigenvectors with small eigenvalues ($i \geq N_\mathrm{I}' + 1, j \geq N_\mathrm{II}' + 1$):

$$\begin{aligned} &\{\varphi_{i'}\}\text{:}\quad i' > N_\mathrm{I}',\quad \sqrt{\overline{|\langle\varphi_{i'}|T^\mathrm{II}\rangle|^2}} \geq \sqrt{\overline{|\langle\varphi_{N_\mathrm{I}'}|T^\mathrm{I}\rangle|^2}};\\ &\{\psi_{j'}\}\text{:}\quad j' > N_\mathrm{II}',\quad \sqrt{\overline{|\langle\psi_{j'}|T^\mathrm{I}\rangle|^2}} \geq \sqrt{\overline{|\langle\psi_{N_\mathrm{II}'}|T^\mathrm{II}\rangle|^2}}. \end{aligned} \tag{14}$$

In other words, only the vectors $\{\varphi_{i'}\}$ should be added that are noticeably collinear with some vectors $\{\psi_j\}_{j=1,\ldots,N_\mathrm{II}'}$ with large eigenvalues and, therefore, can describe (together with $\{\varphi_i\}_{i=1,\ldots,N_\mathrm{I}'}$) structure type II. Similar considerations apply to $\{\psi_{j'}\}$. For the spectral approach, selection rule (14) takes the following form (see Eqs. (11) and (13) at $m = n$):

$$\begin{aligned} &\{\varphi_{i'}\}\text{:}\quad i' > N_\mathrm{I}',\quad \sqrt{\sum_{j=1}^{N^2} \mu_j |\langle\varphi_{i'}|\psi_j\rangle|^2} \geq \sqrt{\lambda_\mathrm{min}};\\ &\{\psi_{j'}\}\text{:}\quad j' > N_\mathrm{II}',\quad \sqrt{\sum_{i=1}^{N^2} \lambda_i |\langle\psi_{j'}|\varphi_i\rangle|^2} \geq \sqrt{\mu_\mathrm{min}}. \end{aligned}$$

Thus, both augmented bases $\{\varphi_i\}_{i=1,\ldots,N_\mathrm{I}'} \cup \{\varphi_{i'}\}$ and $\{\psi_j\}_{j=1,\ldots,N_\mathrm{II}'} \cup \{\psi_{j'}\}$ are complete (within a tolerable accuracy) for describing structures of both the proper and the alternative types. In this approach, precisely the small eigenvalues $\lambda_{i'}$ and $\mu_{j'}$ play the decisive role in the classification criterion, because they are involved in the corresponding matrices $(K_{aa}^\mathrm{I})^{-1}$ and $(K_{aa}^\mathrm{II})^{-1}$ of the likelihood ratio with large weights $1/\lambda_{i'}$ and $1/\mu_{j'}$. In this sense, the method of defining $L(\mathbf{r})$ pro-

posed in this paper can be regarded as the general Pisarenko's method [20].

The above optimal approach simultaneously classifies only two structure types in the image. If it is necessary to classify a greater number of types, one should use multialternative decision rules (see Ch. 13 in [21]). Then, all possible alternative structures are specified a priori and a decision about whether or not a particular type of structure is present is made based on the comparison of distributions $\Lambda'(\mathbf{r})$ obtained for all pairs of different structures.

3. MODIFIED BARTLETT'S AND PISARENKO'S METHODS

The general optimal approach yields the particular Bartlett's and Pisarenko's classification methods modified for application to the problem under consideration. Under certain conditions, these methods prove to be rather efficient. In this formulation, by Bartlett's method, we mean the direct use of the image spectrograms without subjecting them to nonlinear processing. The classical Pisarenko's harmonic expansion method, however, relies on an analysis of eigenvalues of the autocovariance matrix; specifically, this method finds one or several smallest eigenvalues and corresponding eigenvectors and processes them. An advantage of these particular methods over the general approach is that they require many less calculations.

The modified Bartlett's method classifies the images in terms of the statistical features that are pronounced in the structure type being classified and, simultaneously, that are hardly observed or absent in the alternative type [20].The process of classification of the structure of the given type occurs with the participation of all those eigenvectors of this type whose presence in the given type of structure is much more pronounced than in the structure of the alternative type. Particularly, to classify a type I structure, this method takes only those vectors $\varphi_{i=i_0}$ of the whole set $\{\varphi_i\}_{i=1,\ldots,N^2}$ that satisfy the following condition: the average statistical value of the expansion coefficient of structure $|T^{\mathrm{I}}\rangle$ for the vector $|\varphi_{i_0}\rangle$, which equals $\sqrt{\overline{|\langle\varphi_{i_0}|T^{\mathrm{I}}\rangle|^2}}$ (see (6)), is greater than a similar value of the expansion coefficient of the structure $|T^{\mathrm{II}}\rangle$ for the same vector $|\varphi_{i_0}\rangle$, which equals $\sqrt{\overline{|\langle\varphi_{i_0}|T^{\mathrm{II}}\rangle|^2}}$:

$$\left[\sqrt{\overline{|\langle\varphi_{i_0}|T^{\mathrm{I}}\rangle|^2}} \equiv \sqrt{\overline{|a^{\mathrm{I}}_{i_0}|^2}}\right] \gg \left[\sqrt{\overline{|\langle\varphi_{i_0}|T^{\mathrm{II}}\rangle|^2}} \equiv \sqrt{\overline{|b^{\mathrm{II}}_{i_0}|^2}}\right]. \quad (15)$$

This method is in essence a spectral one, because selection criterion (15) does not allow for the off-diagonal elements $\overline{a^{\mathrm{I}}_m (a^{\mathrm{I}}_n)^*}$, $n \neq m$, which become nonzero when morphological classification features are taken into account and thereby carry the main information on the morphology of the structure (see Section 4 below). Then, with Eqs. (11) and (13), relationship (15) takes the form

$$\sqrt{\lambda_{i_0}} \gg \sqrt{\sum_{j=1}^{N^2} \mu_j |\langle\varphi_{i_0}|\psi_j\rangle|^2},$$
$$\text{or} \quad h_0\sqrt{\lambda_{i_0}} \geq \sqrt{\sum_{j=1}^{N^2} \mu_j |\langle\varphi_{i_0}|\psi_j\rangle|^2}, \quad (16)$$

where the coefficient $h_0 < 1$ defined a priori is determined by the costs and probabilities of the errors.

Similarly to formulas (15) and (16), the classification of structure type II uses only such eigenvectors $\psi_{j=j_0}$ of the complete set $\{\psi_j\}_{j=1,\ldots,N^2}$ that satisfy the conditions

$$\left[\sqrt{\overline{|\langle\psi_{j_0}|T^{\mathrm{II}}\rangle|^2}} \equiv \sqrt{\overline{|a^{\mathrm{II}}_{j_0}|^2}}\right] \gg \left[\sqrt{\overline{|\langle\psi_{j_0}|T^{\mathrm{I}}\rangle|^2}} \equiv \sqrt{\overline{|b^{\mathrm{I}}_{j_0}|^2}}\right];$$

i.e., $\sqrt{\mu_{j_0}} \gg \sqrt{\sum_{i=1}^{N^2} \lambda_i |\langle\psi_{j_0}|\varphi_i\rangle|^2}$ or $h_0\sqrt{\mu_{j_0}} \geq \sqrt{\sum_{i=1}^{N^2} \lambda_i |\langle\psi_{j_0}|\varphi_i\rangle|^2}$ $(h_0 < 1)$.

Thus, in fact, the modified Bartlett's method classifies a structure of the given type by using the eigenvectors of the same type that correspond to relatively large eigenvalues. The resultant responses $R^{\mathrm{I}}_{\mathrm{Bt}}(\mathbf{r})$ and $R^{\mathrm{II}}_{\mathrm{Bt}}(\mathbf{r})$ to the zero-mean image $\Gamma_{\mathrm{dif}}(\mathbf{r})$ being classified appear to be strong in regions with structure type I or II, respectively, and are weak in regions with the alternative structure type:

$$R^{\mathrm{I}}_{\mathrm{Bt}}(\mathbf{r}) \equiv \frac{1}{N^{\mathrm{I}}_{\mathrm{Bt}}} \sum_{i_0} \left|F^{\mathrm{I}}_{i_0}(\mathbf{r})\right|,$$

$$R^{\mathrm{II}}_{\mathrm{Bt}}(\mathbf{r}) \equiv \frac{1}{N^{\mathrm{II}}_{\mathrm{Bt}}} \sum_{j_0} \left|F^{\mathrm{II}}_{j_0}(\mathbf{r})\right|,$$

$$F^{\mathrm{I}}_{i_0}(\mathbf{r}) \equiv \int \Gamma_{\mathrm{dif}}(\mathbf{r}')\varphi_{i_0}(\mathbf{r}'-\mathbf{r})d\mathbf{r}',$$
$$F^{\mathrm{II}}_{j0}(\mathbf{r}) \equiv \int \Gamma_{\mathrm{dif}}(\mathbf{r}')\psi_{j_0}(\mathbf{r}'-\mathbf{r})d\mathbf{r}',$$

where $N^{\mathrm{I}}_{\mathrm{Bt}}$ and $N^{\mathrm{II}}_{\mathrm{Bt}}$ are the total numbers of vectors φ_{i_0} and ψ_{j_0}, respectively. Replacing $N^{\mathrm{I}}_{\mathrm{Bt}}$ and $N^{\mathrm{II}}_{\mathrm{Bt}}$ with the sums $\sum_{i_0}\sqrt{\overline{|\langle\varphi_{i_0}|T^{\mathrm{I}}\rangle|^2}} = \sum_{i_0}\sqrt{\lambda_{i_0}}$ and $\sum_{j_0}\sqrt{\overline{|\langle\psi_{j_0}|T^{\mathrm{II}}\rangle|^2}} = \sum_{j_0}\sqrt{\mu_{j_0}}$, one obtains a more accurate normalization.

Another version of our classification process selects the eigenvectors of a specified type that directly produce a weak response of the image being classified in

regions whose structure is of the given type. In this sense, this method can be referred to as the modified Pisarenko's method. In contrast to Bartlett's method, it classifies the structures in terms of a statistical feature that is weakly observed or absent in the structure type being classified and, simultaneously, is pronounced in the alternative type of structure [20]. The criterion for selecting eigenvectors $\varphi_{i=i_0}$ used to classify structure type I is opposite to criterion (15): $\sqrt{\overline{|\langle\varphi_{i_0}|T^{\mathrm{I}}\rangle|^2}} \ll \sqrt{\overline{|\langle\varphi_{i_0}|T^{\mathrm{II}}\rangle|^2}}$; i.e.,

$$\sqrt{\lambda_{i_0}} \ll \sqrt{\sum_{j=1}^{N^2} \mu_j |\langle\varphi_{i_0}|\psi_j\rangle|^2}$$
$$\text{or} \quad H_0\sqrt{\lambda_{i_0}} \le \sqrt{\sum_{j=1}^{N^2} \mu_j |\langle\varphi_{i_0}|\psi_j\rangle|^2}, \tag{17}$$

where $H_0 > 1$. It should be noted that the classical Pisarenko's method analyzes the structure of the given type using eigenvectors of the same type that correspond to the smallest eigenvalues. At the same time, more general criterion (17) does not impose stringent limitations on the magnitude of λ_{i_0}, which may be not too small.

By analogy with formulas (17), to classify structure type II, we select only such vectors $\psi_{j=j_0}$ of the set $\{\psi_j\}_{j=1,\ldots,N^2}$ that $\sqrt{\overline{|\langle\psi_{j_0}|T^{\mathrm{II}}\rangle|^2}} \ll \sqrt{\overline{|\langle\psi_{j_0}|T^{\mathrm{I}}\rangle|^2}}$; i.e.,

$$\sqrt{\mu_{j_0}} \ll \sqrt{\sum_{i=1}^{N^2} \lambda_i |\langle\psi_{j_0}|\varphi_i\rangle|^2}$$
$$\text{or} \quad H_0\sqrt{\mu_{j_0}} \le \sqrt{\sum_{i=1}^{N^2} \lambda_i |\langle\psi_{j_0}|\varphi_i\rangle|^2} \quad (H_0 > 1).$$

Now, the values of $|F^{\mathrm{I}}_{i_0}(\mathbf{r})|$ and $|F^{\mathrm{II}}_{j_0}(\mathbf{r})|$ are minimal in the localization regions of the structure type being sought. To enhance visual perception, it is convenient to ascribe the greatest resultant responses $R^{\mathrm{I}}_{\mathrm{Ps}}(\mathbf{r})$ and $R^{\mathrm{II}}_{\mathrm{Ps}}(\mathbf{r})$ to these regions through the inversion

$$R^{\mathrm{I}}_{\mathrm{Ps}}(\mathbf{r}) \equiv 1/\left[\left\{\frac{1}{N^{\mathrm{I}}_{\mathrm{Ps}}}\sum_{i_0}\left|F^{\mathrm{I}}_{i_0}(\mathbf{r})\right|\right\} + \eta^{\mathrm{I}}_{\mathrm{Ps}}\right],$$

$$R^{\mathrm{II}}_{\mathrm{Ps}}(\mathbf{r}) \equiv 1/\left[\left\{\frac{1}{N^{\mathrm{II}}_{\mathrm{Ps}}}\sum_{j_0}\left|F^{\mathrm{II}}_{j_0}(\mathbf{r})\right|\right\} + \eta^{\mathrm{II}}_{\mathrm{Ps}}\right],$$

where $N^{\mathrm{I}}_{\mathrm{Ps}}$ and $N^{\mathrm{II}}_{\mathrm{Ps}}$ are the total numbers of selected φ_{i_0} and ψ_{j_0}, respectively ($N^{\mathrm{I}}_{\mathrm{Ps}}$ and $N^{\mathrm{II}}_{\mathrm{Ps}}$ may again be replaced by a weighted sum); and $\eta^{\mathrm{I}}_{\mathrm{Ps}}$ and $\eta^{\mathrm{II}}_{\mathrm{Ps}}$ are the regularization coefficients. However, strong responses $R^{\mathrm{I}}_{\mathrm{Ps}}(\mathbf{r})$ and $R^{\mathrm{II}}_{\mathrm{Ps}}(\mathbf{r})$ occur here not only in regions whose structure is of the corresponding type, but also in the regions where $\Gamma_{\mathrm{dif}}(\mathbf{r}) \equiv 0$. Such an undesirable effect is excluded by destructively combining the responses to structures of the two types: at $\eta^{\mathrm{I}}_{\mathrm{Ps}} = \eta^{\mathrm{II}}_{\mathrm{Ps}}$, the difference $R^{\mathrm{I}}(\mathbf{r}) - R^{\mathrm{II}}(\mathbf{r})$ compensates for the response to regions $\Gamma_{\mathrm{dif}}(\mathbf{r}) \equiv 0$, while this difference is positive or negative in regions with structure type I or II, respectively. Therefore, the difference $R^{\mathrm{I}}_{\mathrm{Ps}}(\mathbf{r}) - R^{\mathrm{II}}_{\mathrm{Ps}}(\mathbf{r})$ plays the classification role of the logarithm of the likelihood ratio.

Thus, the possibility of classifying by Bartlett's and Pisarenko's methods depends on the particular set of eigenvectors and corresponding eigenvalues of the types considered. Unlike the general approach based on the likelihood ratio, which always separates the structures of given types in the optimum manner, Bartlett's and Pisarenko's methods separate the structures only if vectors with the contrast properties necessary for classification can be found among all of the eigenvectors. The existence of such vectors at admissible values of h_0 and H_0 may serve as a distinguishability criterion for the given type of structure by the corresponding method.

4. INCORPORATION OF MORPHOLOGICAL FEATURES INTO THE CLASSIFICATION PROCESS

The spatial–spectrum version of the optimal method analyzed in Sections 2 and 3 assumes that the combined matrices are constructed using the complete hypothetical ensemble, which consists of all possible images whose correlation properties are described by given estimate (1) of the autocovariance matrix $A(\mathbf{r}, \mathbf{r}')$, constructed in the form of a block Toeplitz matrix. At the same time, if the process is spatially homogeneous and ergodic, the autocovariance matrix $A(\mathbf{r}, \mathbf{r}')$ is one-to-one related through Wiener–Khintchin's theorem to the spatial power spectrum averaged over the complete ensemble. This spectrum approximately coincides with the spatial spectrum of the sample autocovariance function $K(\boldsymbol{\rho})$ of the given type. The power spectrum, unlike the amplitude spectrum, loses the initial information about the relative phase behavior of different frequency components characteristic of the structure type being classified. As a consequence, $A(\mathbf{r}, \mathbf{r}')$ also does not carry this information. The matrix $A(\mathbf{r}, \mathbf{r}')$ constructed in accordance with relation (1) is a block Toeplitz matrix, but it does not possess the property of cyclicity. Therefore, its eigenvectors (Karhunen–Loéve basis (2), (3))

are not exactly harmonic two-dimensional functions, and an expansion of a chosen image fragment or a structure of a certain type in terms of these functions is not a rigorous Fourier transform. However, the elements of matrix $A(\mathbf{r}, \mathbf{r}')$ with the greatest magnitude are concentrated in a narrow band near its main diagonal, and the difference of this matrix from a cyclic matrix is not very significant. Due to this fact, the eigenvectors are close to harmonic functions, and, when relationships (9)–(13) are used, the optimal classification method is close to the spectral method. In this approach, the combined matrices are constructed based on the two Karhunen–Loéve bases. Therefore, the combined matrices actually contain the same information about the structure types being classified as that contained in the power-spectrum properties of the structures, and no other information is present. Thus, we actually deal with a spectral approach in which the classification does not allow for the morphological features of the structure to be sought. This circumstance will be illustrated below by an example of numerical processing of an artificially generated random combined structure.

The loss of the significant information on morphological features in the purely spectral approach means that it does not use all the possibilities that provide for the most reliable result of classification. In this regard, in this section we introduce morphological features into the spectral approach. To this end, we allow for a certain correlation between the expansion coefficients when constructing combined matrices (7). The proper and alternative bases $\{\varphi_i\}$ and $\{\psi_j\}$ remain the same. Hence, the expansion coefficients $|(X_\mathbf{r})^{\mathrm{I}}_{\mathrm{II}}\rangle$ and $|(X_\mathbf{r})^{\mathrm{II}}_{\mathrm{I}}\rangle$ also do not change (see Eqs. (8)). However, now, when we construct correlation matrices of expansion coefficients (6), we replace averaging over the complete hypothetical ensemble (as in the purely spectral approach) with averaging over a subset of this ensemble. This subset, which is also hypothetical, consists of all image realizations of the structure of a certain type with the sought-after morphological features. It has the same average spatial power spectrum as the complete ensemble; i.e., all images of the subset can be expanded into the two Karhunen–Loéve bases, as before. However, the coefficients of this expansion become correlated in a certain manner, which reflects the morphological features of this type of structure.

Our numerical realization of the above spectral–morphological approach represented the hypothetical subset by a particular learning set in the form of samples of type I or II, respectively (q is the sample number): $u_{\mathrm{dif}} = |T^{\mathrm{I}}_q\rangle_{q=1,\ldots,Q_1}$ or $u_{\mathrm{dif}} = |T^{\mathrm{II}}_q\rangle_{q=1,\ldots,Q_2}$. Then, averaging over the complete hypothetical ensemble in Eqs. (6) is replaced with averaging over the selected samples of the corresponding type, and the correlation obtained in this manner is further used in the matrices K_{aa}, K_{ab}, K_{ba}, and K_{bb}:

$$(K^{\mathrm{I}}_{aa})_{mn} \equiv \overline{a^{\mathrm{I}}_m (a^{\mathrm{I}}_n)^*} = \frac{1}{Q_1}\sum_{q=1}^{Q_1} \langle\varphi_m|T^{\mathrm{I}}_q\rangle\langle T^{\mathrm{I}}_q|\varphi_n\rangle,$$

$$(K^{\mathrm{I}}_{bb})_{mn} \equiv \overline{b^{\mathrm{I}}_m (b^{\mathrm{I}}_n)^*} = \frac{1}{Q_1}\sum_{q=1}^{Q_1} \langle\psi_m|T^{\mathrm{I}}_q\rangle\langle T^{\mathrm{I}}_q|\psi_n\rangle,$$

$$(K^{\mathrm{I}}_{ab})_{mn} \equiv \overline{a^{\mathrm{I}}_m (b^{\mathrm{I}}_n)^*} = \frac{1}{Q_1}\sum_{q=1}^{Q_1} \langle\varphi_m|T^{\mathrm{I}}_q\rangle\langle T^{\mathrm{I}}_q|\psi_n\rangle,$$

$$(K^{\mathrm{I}}_{ba})_{mn} = (K^{\mathrm{I}}_{ab})^*_{nm}.$$

The relationships for matrices of type II are similar. Since realizations $\{|T^{\mathrm{I}}_q\rangle_{q=1,\ldots,Q_1}\}$ and $\{|T^{\mathrm{II}}_q\rangle_{q=1,\ldots,Q_2}\}$ are only finite samples of the hypothetical subset, which does not coincide with the complete ensemble, in the spectral–morphological approach, unlike the spectral approach, the matrices

$$\overline{|T^{\mathrm{I}}\rangle\langle T^{\mathrm{I}}|} = \frac{1}{Q_1}\sum_{q=1}^{Q_1} |T^{\mathrm{I}}_q\rangle\langle T^{\mathrm{I}}_q|, \quad \overline{|T^{\mathrm{II}}\rangle\langle T^{\mathrm{II}}|} = \frac{1}{Q_2}\sum_{q=1}^{Q_2} |T^{\mathrm{II}}_q\rangle\langle T^{\mathrm{II}}_q|$$

do not coincide with matrices A^{I} and A^{II} (i.e., Eqs. (9) and (10) are not valid), a fact that introduces the morphological features. In fact, in the spectral approach, the correlation matrices of expansion coefficients are exactly expressed in the form of Eqs. (11)–(13) in terms of the eigenvalues and inner products of eigenvectors belonging to the Karhunen–Loéve bases. This circumstance shows that the classification process takes into account only the spatial-spectrum features. In the spectral–morphological approach, relationships (11)–(13) become invalid. In particular, the correlation matrices K^{I}_{aa} and K^{II}_{aa} of the structure expansion coefficients in terms of the proper basis now lose their diagonal form (Eq. (11)), which means that coefficients a^{I}_m and a^{I}_n and a^{II}_m and a^{II}_n ($n \neq m$) were uncorrelated when the averaging in the spectral approach was performed over the complete hypothetical basis. Thus, the correlation of these coefficients (at least in the sense of their sign) that appears in the spectral–morphological approach allows for not only spectral, but also morphological features of the structure. The inclusion of morphological features thereby narrows the class of structures that will be referred to the given type as a result of classification, as compared to the spectral approach.

5. MULTIPARAMETER NONLOCAL CLASSIFICATION

The spectral and spectral–morphological approaches considered above can be generalized to classification in

several parameters simultaneously. Namely, a decision on whether or not a given complex structure belongs to a certain type is made based on the combined processing of a set of images $\Gamma_{\text{dif}}^{(\beta)}(\mathbf{r})$ of this structure, which are spatial distributions of each of its parameters specified by the index $\beta = 1, \ldots, B$. In particular, in acoustic tomograpy images, these parameters may be the velocity of sound, density, absorption, the nonlinear parameter, and so on. Then, the learning sets must be given for each parameter in the form of samples $T^{\mathrm{I}(\beta)}$ and $T^{\mathrm{II}(\beta)}$ for each of the two structure types I and II to be identified.

One of the possible ways to construct the vectors and matrices that enter into the classification likelihood ratio suggests that vectors $|T^{\mathrm{I}}\rangle$, $|T^{\mathrm{II}}\rangle$, and $|X_{\mathbf{r}}\rangle$ are initially formed by combining the corresponding vectors for different β and have BN^2 components instead of N^2:

$$|T^{\mathrm{I}}\rangle = \begin{pmatrix} |T^{\mathrm{I}(\beta=1)}\rangle \\ \ldots \\ |T^{\mathrm{I}(\beta=B)}\rangle \end{pmatrix}; \quad |T^{\mathrm{II}}\rangle = \begin{pmatrix} |T^{\mathrm{II}(\beta=1)}\rangle \\ \ldots \\ |T^{\mathrm{II}(\beta=B)}\rangle \end{pmatrix};$$

$$|X_{\mathbf{r}}\rangle = \begin{pmatrix} |X_{\mathbf{r}}^{(\beta=1)}\rangle \\ \ldots \\ |X_{\mathbf{r}}^{(\beta=B)}\rangle \end{pmatrix}.$$

Then, the autocovariance matrices A^{I} and A^{II} and expansion coefficients are formally constructed based on the same relationships as in the one-parameter case. However, in this version, A^{I} and A^{II} are block matrices; each block $(A^{\mathrm{I}})^{\beta 1, \beta 2}$ and $(A^{\mathrm{II}})^{\beta 1, \beta 2}$ is specified by a pair of indices ($\beta 1$, $\beta 2$), where $\beta 1, \beta 2 = 1, \ldots, B$. Matrices A^{I} and A^{II}, whose size is $BN^2 \times BN^2$, have BN^2 eigenvectors and eigenvalues: $\{\varphi_i\}$, $\{\lambda_i\}$ for A^{I} and $\{\psi_j\}$, $\{\mu_j\}$ for A^{II}, where $i, j = 1, \ldots, BN^2$. The expansion coefficients c_i^{I}, c_j^{II}, a_i^{I}, a_j^{II}, b_i^{I}, and b_j^{II} are formally calculated from the same relationships (4) and (5), and their correlation matrices K_{aa}, K_{ab}, and K_{bb}, from Eqs. (6), where $m, n = 1, \ldots, BN^2$. As before, in the purely spectral approach, we have $\overline{|T\rangle\langle T|} = A$; in the spectral–morphological approach, we have $\overline{|T\rangle\langle T|} = \frac{1}{Q}\sum_{q=1}^{Q}|T_q\rangle\langle T_q|$ (Q is the number of realizations in the learning set) for the corresponding structure type, but these matrices $\overline{|T\rangle\langle T|}$ consist of blocks $\overline{|T^{(\beta 1)}\rangle\langle T^{(\beta 2)}|}$.

Local classification methods decide the structure type by comparing one or several parameters of this structure at a particular point of the coordinate or spectral space with initial given values. By contrast, the method proposed in this paper is nonlocal and its multiparameter realization requires knowledge of the correlation matrices $\overline{|T^{\mathrm{I}(\beta 1)}\rangle\langle T^{\mathrm{I}(\beta 2)}|}$ and $\overline{|T^{\mathrm{II}(\beta 1)}\rangle\langle T^{\mathrm{II}(\beta 2)}|}$, which account for the autocorrelation properties of the given type of structure at a fixed parameter $\beta 1 = \beta 2$, as well as for cross-correlation properties at different parameters $\beta 1 \neq \beta 2$, including the multiparameter morphological characteristics.

6. TEST ILLUSTRATION OF THE DIFFERENCE BETWEEN THE SPATIAL–SPECTRUM AND SPECTRAL–MORPHOLOGICAL APPROACHES

As we mentioned above, the spatial–spectrum approach classifies images based on the information on the spatial power spectrum of the sought-after brightness structure type. To clearly illustrate this circumstance, we considered a model Gaussian random structure whose spatial spectrum, with equal initial levels, contains all spatial frequencies observed in the spectrum of an acoustic microscopic image of mammary-gland tissue (Fig. 1b). The two chosen types of this tissue were the large-cluster and small-grain tissues, whose typical regions are outlined in Fig. 1b by the white and black contours, respectively. The spatial spectrum components of the random structure were multiplied by respective root-mean-square values of the spectrum of the large-cluster (in the first version) or small-grain (in the second version) regions. Two test structures were thus synthesized, whose spatial power spectra were tailored to fit the power spectra of each of the two above types of real biological tissue. The brightness variances of these random structures with corrected spectra were additionally made equal to average spatial variances of real tissues of the types considered. Next, we constructed a combined image in the form of a collage. Two of its quadrants were filled with the type I structure described above; two other quadrants, with structure type II (Fig. 1a). Note that, even at a glance, these images and the corresponding real structure types are morphologically dissimilar to the corresponding real tissue types (see Figs. 1a and 1b). The processing of this collage by the spectral method using the bases corresponding to the two types of the mammary-gland tissue classifies the quadrants as large-cluster or small-grain tissue: Fig. 1c shows a nonsmoothed classification result normalized by the maximum value of the result of a similar classification of a real tissue image. (We intend to present a detailed description of the classification results for a real mammary-gland tissue image in our next paper devoted to numerical image processing.) Thus, the purely spectral approach does not discriminate between morphologically different structures whose spectra are statistically equal. This approach does not use the information on the particular shape of characteristic features of the structure being sought, because it ignores the fact that completely different structures can have the same spatial power spectrum. The spectral–morphological approach removes

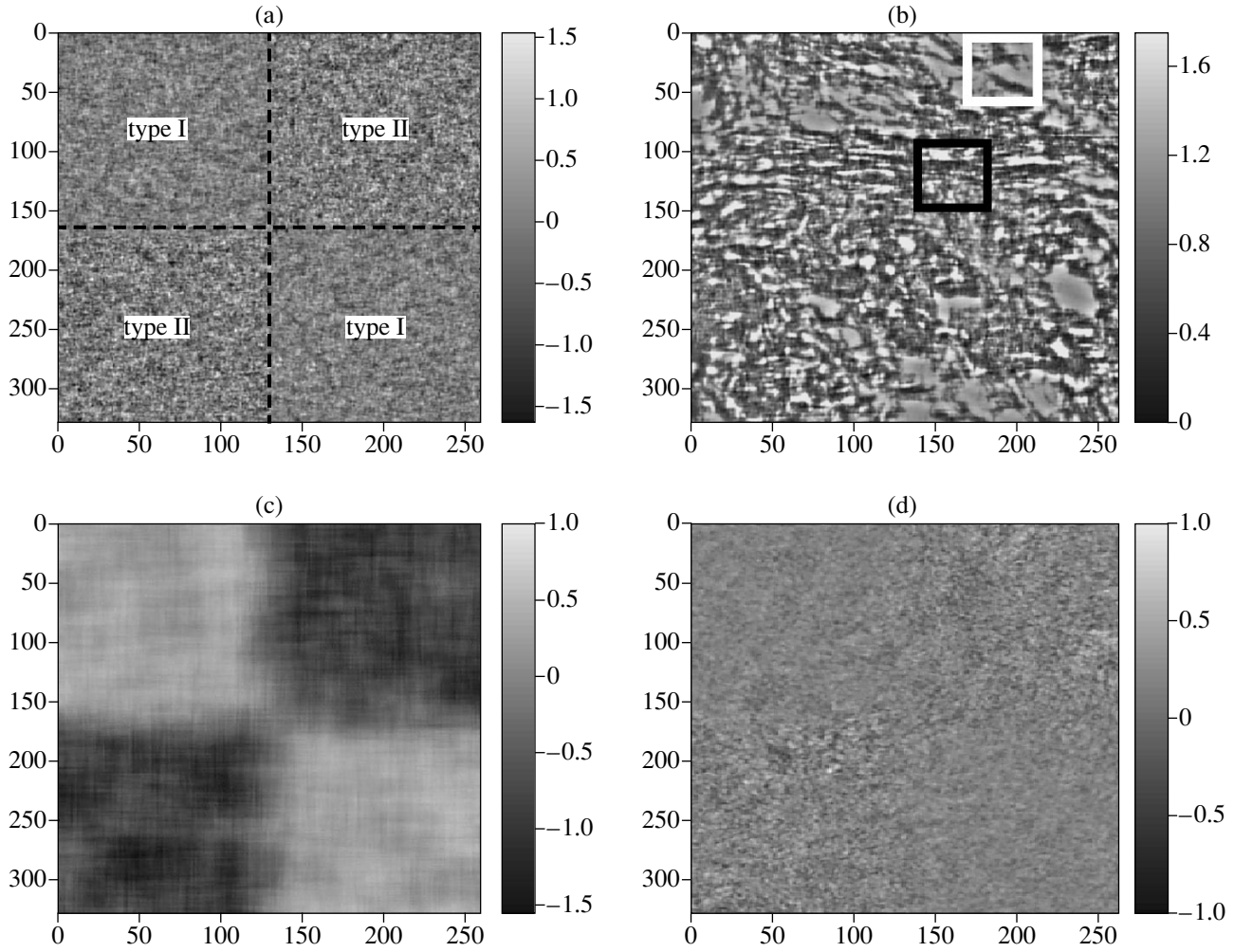


Fig. 1. Illustration of the disadvantage of classification in terms of only the spectrum features and the removal of this disadvantage with allowance for the morphological features of the structure: (a) combined image of two types of random structures whose spatial power spectrum is proportional to the average power spectrum of (b) the large-cluster and small-grain mammary-gland tissue (it is seen that the random structure and the structure of the mammary gland are morphologically dissimilar, although their spatial spectra are close); the results of classification of this combined image with the (c) spectral and (d) spectral–morphological methods.

this disadvantage. Classification with its help does not ascribe the respective quadrants of the collage to any of the given types (Fig. 1d).

7. CONCLUSIONS

The result of the above study is that the Bayesian method was used to describe several seemingly unrelated approaches from a single viewpoint and to find the regions where it is reasonable to apply them. In our next paper, we intend to report on the numerical results obtained with the general and particular methods and to discuss them with examples of the classification of acoustomicroscopic images of real biological tissues and composite structures. These results may be used as an illustration in a comparative analysis of the methods as applied to practical problems.

ACKNOWLEDGMENTS

This work was supported by a grant from the President of the Russian Federation (project no. NSh-1575.003.2) and by the Russian Foundation for Basic Research (project no. 04-02-16043).

REFERENCES

1. *Methods of Computer Processing of Images*, Ed. by V. A. Soĭfer, 2nd ed. (Fizmatlit, Moscow, 2003) [in Russian].
2. J. S. Bleck, M. Gebel, R. Hebel, *et al.*, Ultrasound Med. Biol. **20**, 521 (1994).
3. R. J. Collaris and A. P. G. Hoeks, Acoust. Imaging **22**, 257 (1996).
4. X. Zhang, S. L. Broschat, and P. J. Flynn, J. Acoust. Soc. Am. **111**, 457 (2002).

5. K.-V. Jenderka, M. Schultz, T. Gaertner, *et al.*, in *Proceedings of IEEE Ultrasonics Symposium* (1995), p. 1181.
6. T. Gaertner, K.-V. Jenderka, H. Schneider, and H. Heynemann, Acoust. Imaging **22**, 365 (1996).
7. L. Landini, F. Santarelli, *et al.*, Acoust. Imaging **19**, 387 (1992).
8. N. A. H. K. Rao and M. Helguera, in *Proceedings of IEEE Ultrasonics Symposium* (1997), p. 1193.
9. L. Landini and L. Verrazzani, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**, 448 (1990).
10. J. M. C. Gallet and J. P. Jones, *Presented at 6th Meeting of the World Federation of Ultrasound in Medicine and Biology* (Copenhagen, 1991).
11. V. A. Dumane, P. M. Shankar, C. W. Piccoli, *et al.*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **49**, 664 (2002).
12. M. P. André, M. Galperin, L. K. Olson, *et al.*, Acoust. Imaging **26**, 453 (2002).
13. M. P. André, M. Galperin, L. K. Olson, and C. Green, in *Abstracts of 27th International Symposium on Acoustical Imaging* (Saarrbrücken, Germany, 2003), p. 68.
14. L. Kahl, R. Orglmeister, and K. J. G. Schmailzl, in *Proceedings of Ultrasonics Symposium* (1997), p. 1173.
15. S. V. Il'in and M. N. Rychagov, in *Proceedings of XIII Session of the Russian Acoustical Society* (GEOS, Moscow, 2001), Vol. 2, p. 214.
16. Yu. P. Pyt'ev, Dokl. Akad. Nauk SSSR **269**, 1061 (1983) [Sov. Phys. Dokl. **28**, 308 (1983)]; Pattern Recogn. Image Anal. **3** (1), 19 (1993).
17. E. A. Grachev, D. M. Ustinin, and A. I. Chulichkov, Mat. Model. **15** (5), 47 (2003).
18. A. I. Chulichkov, E. A. Grachev, D. M. Ustinin, and E. A. Cheremukhin, Microelectron. Eng. **69**, 555 (2003).
19. V. A. Burov, O. V. Dmitriev, and I. P. Polyakova, Radiotekh. Élektron. (Moscow), No. 11, 2285 (1974).
20. S. L. Marple, Jr., *Digital Spectral Analysis with Application* (Prentice-Hall, Englewood Cliffs, N.J., 1987; Mir, Moscow, 1990).
21. B. R. Levin, *Theoretical Fundamentals of Statistical Radio Engineering* (Radio i Svyaz', Moscow, 1989), p. 656 [in Russian].

Translated by A. Khzmalyan

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 63–70. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 81–89.

Application of Focusing Arrays to the Problems of Acoustic Brightness Thermometry

V. A. Vilkov, E. V. Krotov, A. D. Mansfel'd, and A. M. Reĭman

Applied Physics Institute, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603950 Russia

e-mail: mansfeld@appl.sci-nnov.ru

Received January 22, 2004

Abstract—Theoretical and experimental studies on the localization of heated objects by the methods of acoustic brightness thermometry are carried out. It is demonstrated that, in the case of using a single focusing array, the spatial localization of heated objects depends on the size of the source. One- and two-dimensional tomography of a real heated source is performed by an acoustic thermal tomograph with a focusing array. The results agree well with the data calculated according to the suggested model. The applicability of correlation focusing acoustic brightness thermometry to the localization of a heated source is investigated both theoretically and experimentally. It is demonstrated that a considerable increase in the spatial resolution of the method leads to a significant loss in sensitivity. © *2005 Pleiades Publishing, Inc.*

The distribution of internal temperature of a human body may contain ample information on the body's state, functioning, and reactions to various effects. It can also be used for the diagnostics of certain disorders, including oncological ones. Monitoring the internal temperature of biological tissues is also necessary in the case of tissue hyperthermia. One of the promising techniques for investigating internal temperature fields is acoustic brightness thermometry based on the reception of equilibrium acoustic radiation caused by the thermal motion of atoms and molecules of the medium. The radiation intensity is directly proportional to the thermodynamic temperature of the medium, and the radiation itself has a broad spectrum. The method of acoustic brightness thermometry provides an opportunity to perform an in-depth detection of acoustic radiation with a relatively high spatial resolution. This is possible thanks to the reception of ultrasonic radiation in millimeter and submillimeter wavelength range, where the sound attenuation in biological tissues is relatively weak. In this case, the ratio of the reception depth L to the wavelength may reach $L/\lambda \sim 60$ [1].

When discussing the opportunities provided by acoustic brightness thermometry, it is possible to separate two major problems that should be solved in the course of the investigation. First is the problem of localizing the heated sources. This problem can be solved using highly directional scanning arrays and applying reconstruction algorithms. Second is the problem of monitoring the temperature of an object during a long time period. Such a problem may arise, for example, in the process of hyperthermia or engraftment.

Up to now, the theoretical foundation of acoustic brightness thermometry has been considered in [1, 2] and some algorithms for the reconstruction of the images of thermal fields of biological objects have been proposed according to the results of scanning the objects with the array of an acoustic brightness thermal tomograph [3–5]. Measurements of the acoustic brightness temperature of a human hand have been conducted [6]. Correlation properties of thermal acoustic radiation have been studied [7–9], and methods for improving the resolution of correlation acoustic brightness thermal tomographs have been proposed [10]. Methods of matching the sensors and improving the efficiency were investigated in [11]. Designs of acoustic brightness thermal tomographs were optimized [12], and multichannel localization of heated objects has been performed experimentally [13]. Ways of applying the methods of acoustic brightness thermal tomography to monitoring the internal temperature of biological objects in the process of laser hyperthermia were studied in [14, 15].

The purpose of this work is to investigate theoretically and experimentally the applicability of focusing arrays, which can be very important for both the problem of localization of heated objects and the problem of continuous monitoring of the internal temperature of an object.

The use of short-focus scanning arrays with large apertures in acoustic brightness thermal tomography provides an opportunity to obtain additional possibilities for the localization of heated regions within biological tissues in both the transverse coordinate and the depth [16]. The cross-sectional area of a beam produced by a focusing array changes as a function of the distance to the array and reaches its minimum at the focus. A redistribution of the intensity of the signal received by the array from different spatial points takes place, and the maximum signal arrives from the focus

"waist." This gives grounds to believe that, by using a focusing array, it is possible to localize heated objects by moving the focus waist.

At the same time, these advantages are not as evident as in the case of the use of focusing transmit–receive transducers in active location [17]. First, in the case of active location, a multiplication of the directivity patterns of an array operating in transmission and reception modes occurs. Second, in the case of the reception mode, the intensity of radiation from the focal zone (the useful signal) does not always exceed the intensity of radiation from other regions (noise). Third, the presence of considerable absorption in some cases completely prevents the object localization. Fourth, as we will demonstrate later, the use of a single focusing array even with a large amplification factor does not ensure the localization in depth for objects with transverse dimensions noticeably exceeding the diameter of the focus waist.

Let us consider several situations for different relationships of the heated-object dimensions and the focus waist and then compare the result obtained with the case of the use of a nonfocusing array. The physical quantity measured by an acoustic thermal tomograph is the acoustic brightness temperature, which is equal to the thermodynamic temperature of an ideal blackbody producing the same acoustic radiation flux as that of the object under investigation [1, 2]. In the case of an array with an inhomogeneous distribution of spatial sensitivity, the acoustic brightness temperature T_a can be represented in the form

$$T_a = \int_0^\infty \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(z) E(x, y, z) T(x, y, z) \times \exp\left(\int_0^z -\gamma(l)\, dl \right) dx\, dy\, dz, \quad (1)$$

where $\gamma(l)$ is the acoustic absorption coefficient (in this model case, we take into account only the dependence on the longitudinal coordinate and ignore the frequency dependence for simplicity), $T(x, y, z)$ is the distribution of the thermodynamic temperature of the object, and $E(x, y, z)$ is the intensity distribution of the array field in the case of operation in the transmission mode.

Let us consider an array with a Gaussian distribution of sensitivity. Exactly this array, which is a piezoelectric transducer shaped as a spherical segment, was implemented in our experiments. Measurements of the reception field of such an array, made with the help of a radiator with a diameter of 0.5 mm, which was positioned in the focus waist of the array, demonstrated that it is close to the Gaussian one [18]. This provides an opportunity to use a corresponding approximation in calculating the field intensity.

In this case, $E(x, y, z)$ can be represented as

$$E(x, y, z) = \frac{\exp\left(\dfrac{-(x^2 + y^2)}{R(z)^2} \right)}{\pi R^2(z)}, \quad (2)$$

where $R^2(z, r) = R_0^2 (1 - z/F)^2 + z^2 (kR_0)^2$, R_0 is the radius of the array aperture, k is the wave number, and F is the geometrical focus.

We write down the temperature distribution in the medium in the form of the sum of the constant temperature (background) T_0 and the Gaussian distribution of the temperature of the heated source:

$$T(x, y, z) = T_0 + \Delta T \exp\left(-\left[\frac{(x - x_0)^2}{\sigma_x^2} + \frac{(y - y_0)^2}{\sigma_y^2} + \frac{(z - z_0)^2}{\sigma_z^2} \right] \right), \quad (3)$$

where x_0, y_0, and z_0 are the coordinates of the source center and σ_x, σ_y, and σ_z are the characteristic dimensions of the source (the half-width of the Gaussian function).

Now, let us consider several particular cases of source positioning with respect to the array for measurements with the help of focusing and nonfocusing arrays. Let an array with $R_0 = 25$ mm, $F = 30$ mm, and $\lambda = 1.5$ mm be positioned on the surface of a biological tissue. The source with $\Delta T = 1$ *K* has the dimensions $\sigma_x = \sigma_y = \sigma_z = 5$ mm and is located at the array axis. The background temperature T_0 is assumed to be equal to zero for simplicity. Physically, this can be realized by subtracting the constant level, which corresponds to an equivalent medium without the source with the thermodynamic temperature T_0, from the received acoustic brightness signal. The dependence of acoustic brightness temperature on the depth of the source position is shown in Fig. 1 (curve *1*). The dependence of the acoustic brightness temperature on the coordinate of the source located on the axis of a nonfocusing array with $R = 5$ mm is given for comparison (curve *2* in Fig. 1).

As one can see from Fig. 1, it becomes possible to localize the source by the focusing array when the source falls into the focus. A sharp drop of the curve at a distance of 45 mm in the case of the nonfocusing array can be explained by the diffraction divergence of the beam.

Figure 2 displays the dependence of the acoustic brightness temperature on the distance between the array and the source for different values of the extension of the heated source along the array axis. As one can see from the plots, an increase in the source dimensions leads to growth of the acoustic brightness temperature even in the case when its dimensions noticeably exceed the dimensions of the focus waist.

It is necessary to note that, using a focusing array, it is impossible to localize in depth the objects extended

in the transverse direction, since the change in the signal intensity, as the object approaches the focus, is exactly compensated by the decrease in the cross-sectional area of the beam and, therefore, in the brightness of the object part that falls into the sensitivity zone of the array. Thus, the detected acoustic brightness temperature of an extended object in the case of its motion changes only because of the presence of acoustic absorption and does not depend on the focusing properties of the array. As a result, the value of the acoustic brightness temperature is shown to depend on the volume of the heated object even if its dimensions exceed the focus diameter.

Figure 3 demonstrates the dependence of the acoustic brightness temperature of an infinite heated layer with a thickness of 10 mm and a source with a transverse dimension of 10 mm on the depth of their position in a biological tissue (the array focus is located at a depth of 50 mm).

It is possible to obtain almost equal acoustic brightness temperatures for different shapes of the source. For example, in the case of a heated layer and an object extended in the transverse direction positioned at a distance of 50 mm from the surface, their acoustic brightness temperatures almost coincide in Fig. 3. It is possible to select the shapes and dimensions of two different sources for which, even in the case of different thermodynamic temperatures, the values of the acoustic brightness temperature are equal; however, the shapes of the curves will be different. Therefore, to reconstruct the true distribution of the thermodynamic temperature, it is necessary to scan an object and subsequently solve the inverse problem that takes into account the parameters of the array field. This means that, even for such a simple procedure as monitoring the temperature of a selected tissue region, for example, in the case of hyperthermia, it is necessary to measure the temperature not at one point but in a certain vicinity, which requires the array scanning and the subsequent reconstruction of the temperature distribution with the help of some tomographic algorithm (see, e.g., [13]).

In the case when the focal distance has a value of the order of magnitude of the absorption depth, the position of the peak of acoustic brightness temperature turns out to be shifted with respect to the focus position, and if the focal distance considerably exceeds the absorption depth, the pronounced peak may be absent (Fig. 4). In this case, localization is impossible, but such an array provides an opportunity to equalize the sensitivities at different points along the array axis.

Now, let us consider another version of scanning. Let an array be located not on the tissue surface but within a weakly absorbing contact fluid, and let it move along and across the axis and receive radiation through a sound-transparent window (Fig. 5). This scanning scheme is most appropriate for implementation in practice. An example of the distribution of acoustic bright-

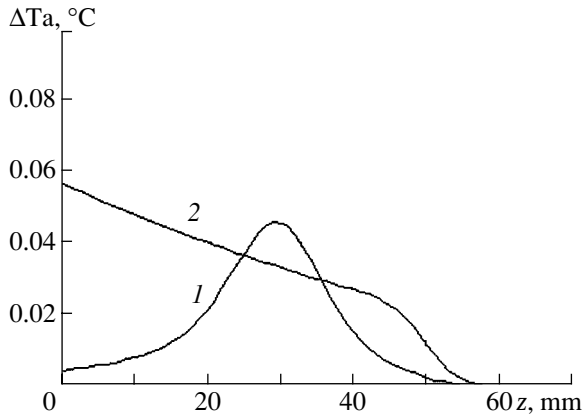


Fig. 1. Dependence of the acoustic brightness temperature on the depth of the heated source position in a biological tissue: measurements with (*1*) a nonfocusing array and (*2*) a focusing array ($\gamma_{medium} = 0.23$ cm^{-1}).

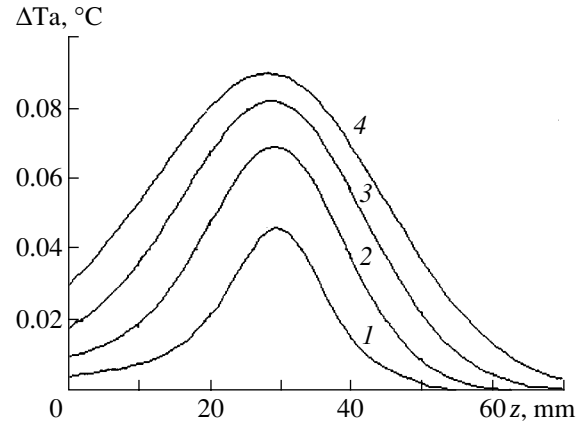


Fig. 2. Longitudinal distribution of the acoustic brightness temperature for different lengths of the source: σ_z = (*1*) 5, (*2*) 10, (*3*) 15, and (*4*) 20 mm. The source is positioned on the array axis ($\gamma_{medium} = 0.23$ cm^{-1}).

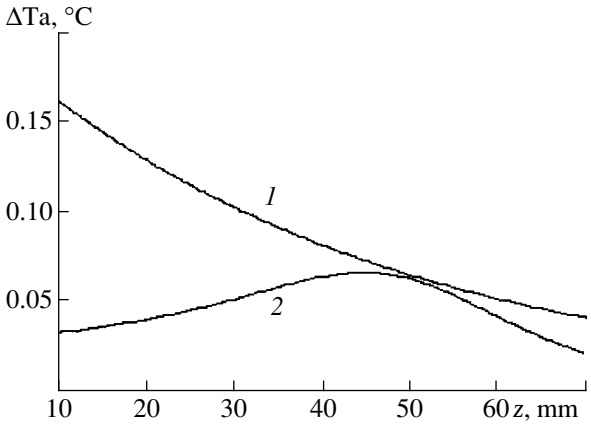


Fig. 3. Dependence of the acoustic brightness temperature of (*1*) a layer and (*2*) a source extended in the transverse direction on the depth of the source position ($\gamma_{medium} = 0.23$ cm^{-1}).

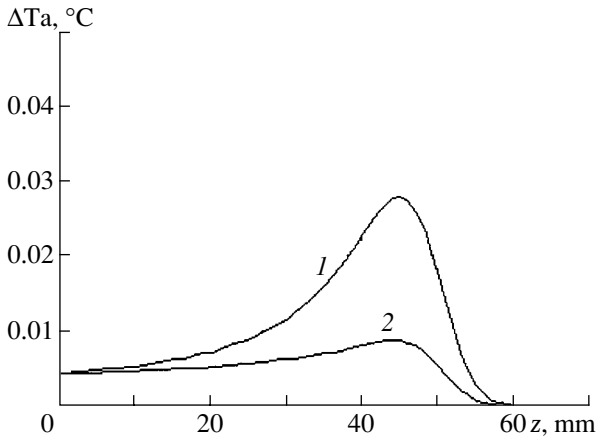


Fig. 4. Dependence of the acoustic brightness temperature on the depth of the source position when the focal distance is (*1*) of the order of magnitude and (*2*) greater than the absorption depth.

ness temperature in the case of scanning a heated object through a layer of a contact fluid is given in Fig. 6.

The experimental verification of the theoretical concepts discussed above was conducted using a thermal tomograph with a focusing array, which had the following parameters: the radius of the array aperture was equal to 30 mm, and the focal distance was $F = 64$ mm. The thermal tomograph was designed according to the compensation scheme [13, 19].

A thin-walled polystyrene tube with a diameter of 10 mm, filled with castor oil and heated by electric current passing through a nichrome wire placed in the tube center, was used as a heated source. It is necessary to note that no direct measurement of acoustic absorption for the oil-filled tube was conducted. The absorption in polystyrene (according to the known reference data) is 0.42 cm^{-1}. As for the oil, the absorption was measured using active ultrasonic location and was equal to 0.33 cm^{-1}. Since the tube was thin-walled (the wall thickness was no greater than 0.5 mm), the absorption in the object was practically determined by the oil.

The sensitivity of the thermal tomograph with calibration according to the model of an ideal blackbody was 0.2 K at a storage time of 5 s, the average operation frequency was 1.5 MHz, and the reception bandwidth was 800 kHz. Figure 7 shows the measured and calculated dependences of acoustic brightness temperature versus the transverse and longitudinal coordinates (the time of a single measurement was 60 s). Figure 8 displays a two-dimensional scan of the distribution of the acoustic brightness temperature.

The longitudinal and transverse localization of a heated source placed under a layer of cow liver with a thickness of 5 mm was also conducted experimentally (Fig. 9). One can see from Fig. 9 that the spatial distributions of the normalized acoustic brightness signal are different for the source in water and within the liver layer, despite the small thickness of the layer. This is connected with the presence of a considerable sound absorption in the layer (the coefficient of acoustic absorption of cow liver at a frequency of 1.5 MHz is $\gamma \approx 0.45$ cm^{-1} [15]).

Thus, we have demonstrated that, first, the localization of small objects whose transverse dimensions are comparable with the dimensions of the focal waist by a scanning thermal tomograph with a focusing array is possible, while it is impossible to localize objects extended in the transverse direction. To reconstruct the temperature profile more accurately, it is necessary to take into account the spatial sensitivity of the array and,

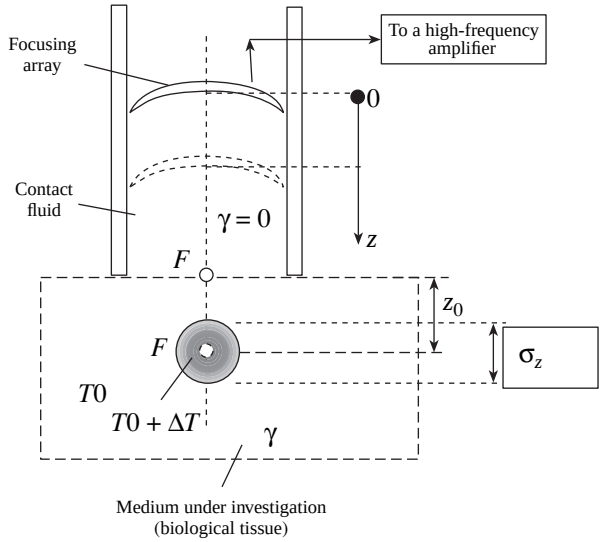


Fig. 5. Schematic diagram of an acoustic brightness thermal tomograph with a focusing array.

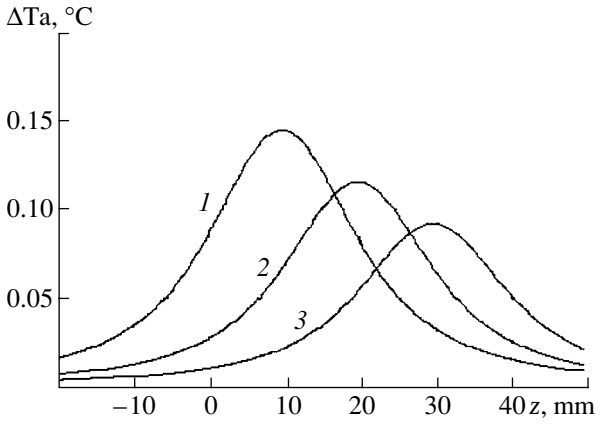


Fig. 6. Variation of the acoustic brightness temperature in the case of the longitudinal scanning for a fixed position of the source at the depth z_0 = (*1*) 10, (*2*) 20, and (*3*) 30 mm below the tissue surface; $F = 40$ and $R_0 = 25$ (γ_{medium} = 0.23 cm^{-1}).

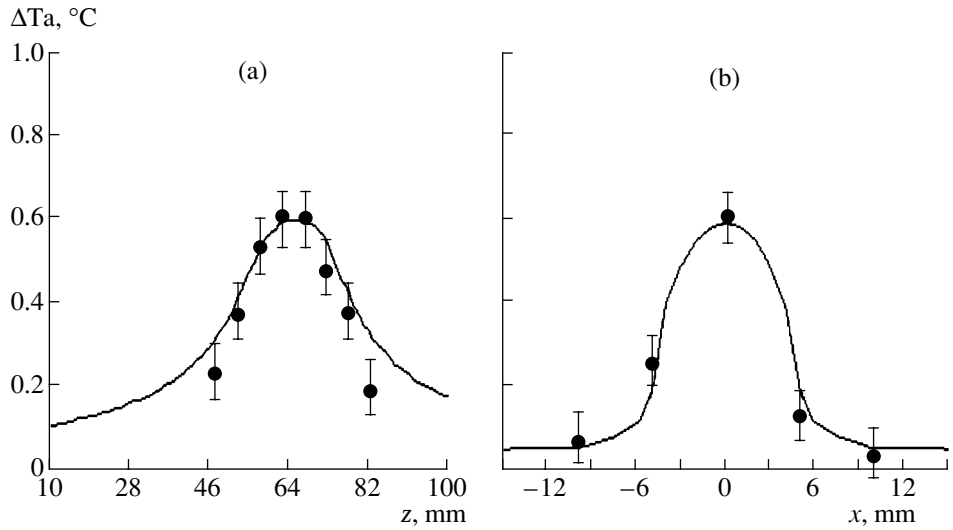


Fig. 7. Experimental data on the localization of a heated source by an acoustic brightness thermal tomograph with a focusing array in different media: (a) result of the longitudinal localization of a heated source positioned in water (an oil-filled tube) and (b) result of the transverse localization of a heated source in water (the dots represent experimental data, and the solid line, the theoretical approximation).

therefore, to solve the inverse problem. Second, despite the high spatial selectivity of a focusing array, the value of the acoustic brightness temperature in the case of a motionless focal waist depends on the dimensions of the source and the temperature distribution in it, which prevents temperature monitoring without scanning the object. Finally, when the values of the focal distance are of the order of magnitude or greater than the absorption depth, it is impossible to localize a heated source.

A further improvement of the localization of heated sources is possible by combining the advantages of focusing arrays and the opportunities provided by correlation signal processing [7–9, 19, 20]. The possibility of using the correlation reception of signals by plane arrays in the problems of acoustic thermal tomography was indicated in [7–9]. However, this method has several drawbacks. The correlation processing of signals can be employed mainly for the localization of point sources of radiation. In the case of distributed sources with dimensions noticeably larger than the spatial correlation radius, the sensitivity of an acoustic thermal tomograph depends on the transverse dimensions of the object, since the signals from different object points arrive at the receiving arrays with different delays. As the band of a received signal is finite and determined by the transmission bandwidth of the array and the receiving amplifier (here and further $\Delta f \approx f_{av}/2$), the correlation function is alternating. As a result of averaging over all object points covered by the receiving beams, the integral value of the response turns out to be close to zero. In particular, it was indicated in [20, 21] that a signal from a distributed source turns out to be nonzero only in the presence of the temperature gradient detected by different channels of the correlation thermal tomograph. However, as will be shown below, the combination of correlation processing and focusing arrays provides an opportunity to obtain a uniquely high spatial resolution. For this purpose, it is necessary to select the appropriate parameters of the correlation function (the spatial correlation length) and the focal waists (the array dimensions, the focal length, and the wavelength).

Technologically, a correlation acoustic thermograph with a focusing array can be designed in several ways. For example, a focusing array is divided into two equal parts, the signals of which are amplified and multiplied, and the result is integrated. We used this array design in the experiments.

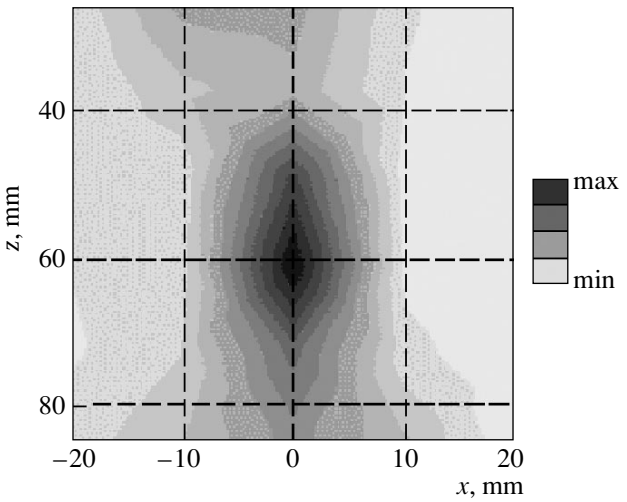


Fig. 8. Two-dimensional scan of a heated source in water.

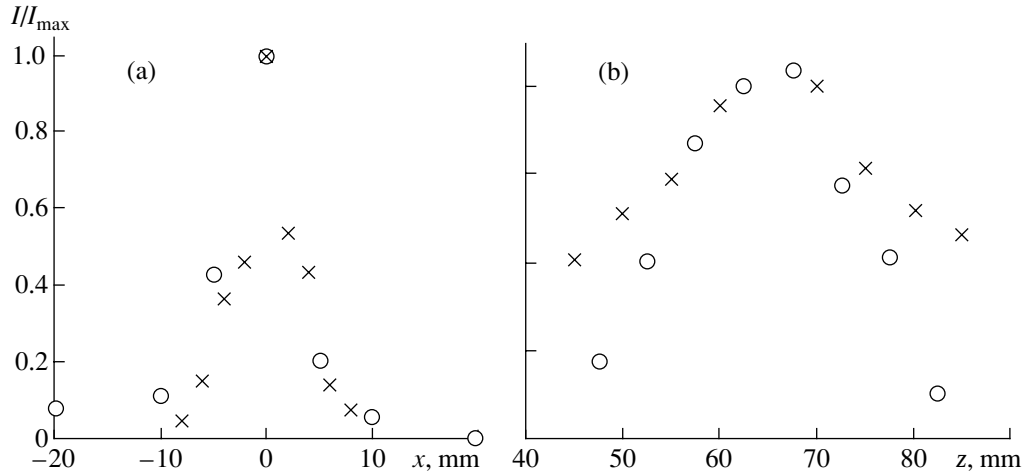


Fig. 9. Experimental results on the localization of a heated source by an acoustic brightness thermal tomograph with a focusing array in different media: (a) result of the transverse localization of a heated source and (b) result of the longitudinal localization of a heated source; the open circles refer to a heated source in water, and the crosses, to a heated source under a layer of cow liver.

Another method is the employment of two focusing arrays positioned at a certain "base" distance from each other. Since there are no fundamental differences between these two methods but theoretical calculations are much simplified, we used the second method to simulate the object localization. In this case, the z axis is directed into the region under investigation and the x and y axes are directed along the axis connecting the array centers and normally to it, respectively. For this array configuration, the acoustic brightness temperature can be represented in the form

$$T_a = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty}\gamma(z)K(x,y,z)E_{12}(x,y,z)T(x,y,z) \times \exp\left(-\int_0^{\infty}\gamma(z)dz\right)dxdydz, \quad (4)$$

where $E_{12}(x, y, z)$ is the product of the spread functions of the array and $K(x, y, z)$ is the cross-correlation function of signals at the output of the thermal tomograph, after a multiplier and an integrator; this function is determined by the frequency bands of the arrays and the radio channel. In the case of a rectangular frequency characteristic of the receiver with an average frequency ω_0 and a half-width $\Delta\omega$, the correlation function has the form

$$K(x,y,z) = \frac{\sin\left(\frac{\Delta\omega\Delta r(x,y,z)}{c}\right)}{\frac{\Delta\omega\Delta r(x,y,z)}{c}}\cos\left(\frac{\omega_0\Delta r(x,y,z)}{c}\right), \quad (5)$$

where $\Delta r(x, y, z) \approx \sqrt{(x+x_0)^2+z^2} - \sqrt{(x-x_0)^2+z^2}$ and x_0 is the coordinate of the array center. We assume that the array centers lie in the xz plane. Two situations can be realized depending on the relationship between the parameters of the correlation function of the signal and the parameters of the spread functions of the arrays. If the width of the spread function considerably exceeds the width of the principal lobe of the correlation function, the average response value is close to zero (Fig. 10a).

When the width of the spread function $E_{12}(x, y, z)$ is considerably narrower than the principal lobe of the correlation function, it determines the spatial resolution (Fig. 10b). The value of the acoustic brightness temperature in this case is also very small because of the small value of the volume from which the signal is received. Thus, in any case, the acoustic brightness signal is small, although the reasons for this are different.

Nevertheless, the last case deserves our attention. Using this tomograph it is possible, first, to localize a layer with a temperature different from the temperature of the surrounding medium (Fig. 11a) and, second, to localize small sources, in which case the value of the acoustic brightness temperature almost does not depend on the source dimensions in a wide range (Fig. 11b). This can be used for monitoring the temperature of a selected region of the medium without scanning the object.

A correlation acoustic brightness thermal tomograph with a focusing array was designed to test experimentally the possibility of the localization of heated objects. The average operation frequency of the acoustic brightness thermal tomograph was 1.5 MHz, and its reception band was 800 kHz. The array, shaped as a spherical segment, was divided in two halves, the signal from which was multiplied and averaged after amplification. The same thin-wall polystyrene tube 10 mm in

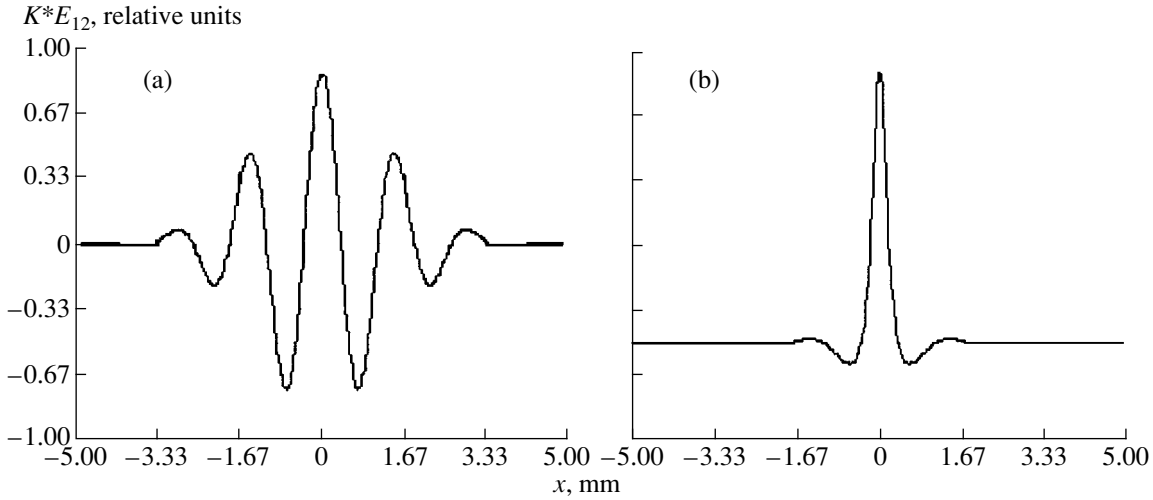


Fig. 10. Dependence of the product of the correlation and spread functions of the arrays (in relative units) on the transverse coordinate (a) when the size of the cross-section zone is greater than the width of the principal lobe of the correlation function and (b) when the size of the cross-section zone is smaller than the width of the principal lobe of the correlation function.

diameter, filled with castor oil, was used as a heated source, as in the experiments described above. The experimentally measured sensitivity of the correlation acoustic brightness thermal tomograph calibrated with respect to a heated source was about 5 K at a storage time of 5 s. It is necessary to note that the problem of sensitivity determination by calibration according to the model of an ideal blackbody is rather difficult in our case and is not considered here. Figure 12 demonstrates the measured dependences of the acoustic brightness temperature (in relative units) on the transverse and longitudinal coordinates. This result shows that, in the case of transverse localization of heated objects, the effect of the negative components of the spatial correlation function at the given characteristics of the acoustic brightness thermal tomograph is substantial, which influences the sensitivity of the method.

Thus, we have conducted theoretical and experimental studies on the localization of heated objects by the methods of focusing acoustic brightness thermometry. It was demonstrated that, in the case of employment of a single focusing array, the value of the measured acoustic brightness temperature depends on the dimensions of the source. The proposed model agrees well with experimental data. Two-dimensional tomography of a heated source by a thermal tomograph with a focusing array was conducted, and theoretical and experimental investigations of the applicability of the correlation focusing acoustic brightness thermometry were performed. As one can see from the curves obtained for the distribution of the acoustic brightness temperature, when the thermodynamic temperature of the source changes by one degree, the acoustic brightness temperature changes by only several tenths of degree in the case of measurement with a focusing array, and by several hundredths of degree in the case

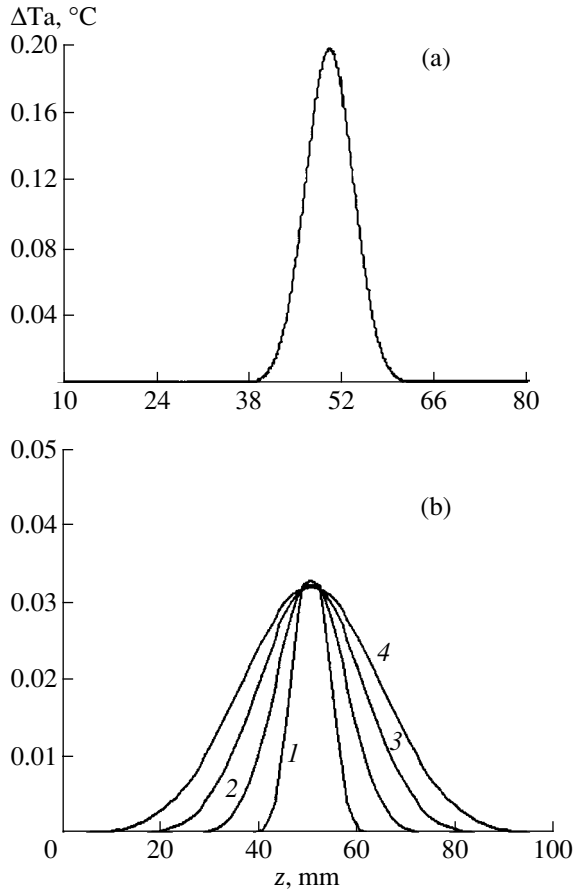


Fig. 11. (a) Example of the localization of a "temperature" layer ($\Delta T = 1$°C, the layer thickness is 10 mm, and the depth of the layer position is 50 mm) by a correlation acoustic brightness thermal tomograph. (b) Variation of the acoustic brightness temperature in the case of longitudinal scanning for different dimensions of a spherical heated source: $\sigma_x = \sigma_y = \sigma_z$ = (*1*) 5, (*2*) 10, (*3*) 15, and (*4*) 20 mm (γ_{medium} = 0.23 cm^{-1}).

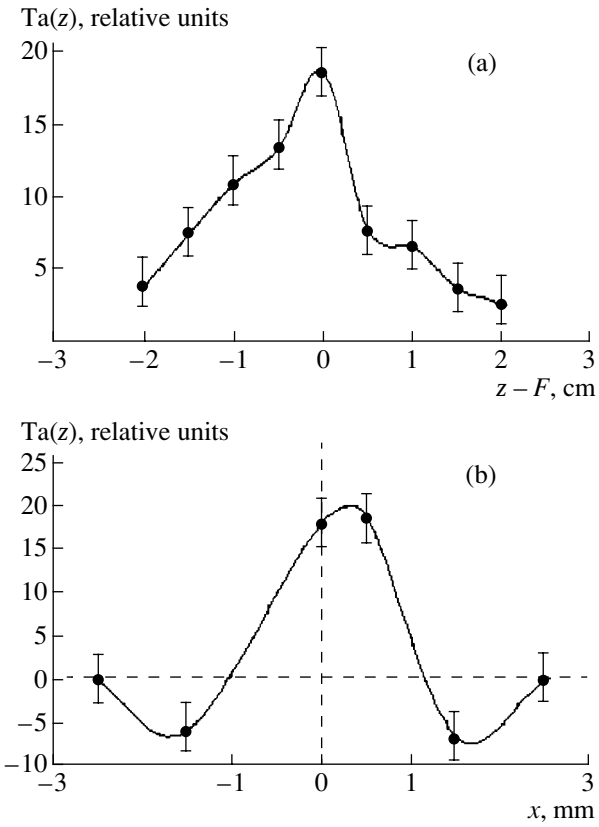


Fig. 12. Experimental results on the longitudinal and transverse localization of a heated source by a correlation acoustic brightness thermal tomograph with a focusing array in water: (a) distribution of the acoustic brightness temperature (in relative units) along the principal axis of the focusing array (the point of the geometrical focus is assumed to be the origin of the *z* axis); (b) distribution of the acoustic brightness temperature (in relative units) across the principal axis of the focusing array at the array focus, normal to the line of the array separation. The dots refer to experimental data and the solid line, to the approximation (γ_{medium} = 0.23 cm^{-1}).

of measurement with a correlation reception, which is caused by the smallness of the region from which the detected signal arrives. This region can be expanded by increasing the wavelength and retaining a constant value of the amplification coefficient of the focusing array. In this case, a decrease in the frequency-dependent absorption coefficient leads to a decrease in the acoustic brightness of the source and also to a decrease in the ultrasound propagation loss along the reception path.

ACKNOWLEDGMENTS

We are grateful to L.S. Dolin for valuable discussions and assistance in theoretical studies. The work was supported by the Russian Foundation for Basic Research, project nos. 01-02-17645 and 03-02-06380.

REFERENCES

1. Yu. N. Barabanenkov and V. I. Pasechnik, Akust. Zh. **40**, 542 (1994) [Acoust. Phys. **40**, 480 (1994)].
2. V. I. Pasechnik, Akust. Zh. **36**, 718 (1990) [Sov. Phys. Acoust. **36**, 403 (1990)].
3. Yu. V. Gulyaev, K. M. Bograchev, I. P. Borovikov, *et al.*, Radiotekh. Élektron. (Moscow) **43** (9), 140 (1998).
4. V. A. Burov and E. E. Kasatkina, Akust. Zh. **43**, 162 (1997) [Acoust. Phys. **43**, 135 (1997)].
5. S. Yu. Ksenofontov, A. D. Mansfel'd, and A. M. Reĭman, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **40**, 752 (1997).
6. A. A. Anosov, K. M. Bograchev, and V. I. Pasechnik, Akust. Zh. **44**, 299 (1998) [Acoust. Phys. **44**, 248 (1998)].
7. Yu. N. Barabanenkov and V. I. Pasechnik, Akust. Zh. **41**, 563 (1995) [Acoust. Phys. **41**, 494 (1995)].
8. A. A. Anosov, M. A. Antonov, and V. I. Pasechnik, Akust. Zh. **46**, 28 (2000) [Acoust. Phys. **46**, 21 (2000)].
9. A. A. Anosov and V. I. Pasechnik, Akust. Zh. **49**, 161 (2003) [Acoust. Phys. **49**, 129 (2003)].
10. V. A. Burov, E. E. Kasatkina, O. D. Rumyantseva, and S. A. Filimonov, in *Proceedings of XI Session of the Russian Acoustical Society* (GEOS, Moscow, 2001), Vol. 2, p. 30.
11. A. G. Sanin, P. K. Chichagov, and A. M. Reĭman, in *Ultrasonic Diagnostics* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1983), pp. 21–36 [in Russian].
12. A. V. Balandin, A. D. Mansfel'd, and A. V. Shishkov, in *Proceedings of XI All-Union Acoustical Conference* (Moscow, 1991), Sect. O, p. 40.
13. E. V. Krotov, S. Yu. Ksenofontov, A. D. Mansfel'd, *et al.*, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **42**, 479 (1999).
14. V. I. Mirgorodsky and A. I. Chmil', http://www.ire.rssi.ru/acoust/publ/STI.htm.
15. E. V. Krotov, M. V. Zhadobov, A. M. Reyman, *et al.*, Appl. Phys. Lett. **81**, 3918 (2002).
16. E. V. Krotov, V. A. Vilkov, A. D. Mansfel'd, and A. M. Reĭman, in *Methods of Acoustic Diagnostics of Inhomogeneous Media* (Inst. Prikl. Fiz. Ross. Akad. Nauk, Nizhni Novgorod, 2002), pp. 235–239 [in Russian].
17. I. N. Kanevskiĭ, *Focusing of Sonic and Ultrasonic Waves* (Nauka, Moscow, 1977) [in Russian].
18. E. V. Krotov, A. D. Mansfel'd, A. M. Reĭman, and A. V. Vilkov, in *Proceedings of XI Session of the Russian Acoustical Society* (GEOS, Moscow, 2001), Vol. 3, pp. 165–169.
19. A. G. Kislyakov, V. A. Razin, and N. M. Tseĭtlin, *Introduction to Radio Astronomy* (Fizmatlit, Moscow, 1995), Part 2 [in Russian].
20. A. Mamouni, Y. Leroy, J. C. Van de Velde, and L. Bellarbi, J. Microwave Power **18** (3), 286 (1983).
21. A. A. Anosov, Yu. N. Barabanenkov, and A. G. Sel'skiĭ, Akust. Zh. **49**, 725 (2003) [Acoust. Phys. **49**, 615 (2003)].

Translated by M. Lyamshev

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 71–80. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 90–100.

Statistical Description of Ray Chaos in an Underwater Acoustic Waveguide

A. L. Virovlyansky

Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603950 Russia
e-mail: viro@hydro.appl.sci-nnov.ru

Received August 4, 2003

Abstract—An approximate analytical approach is developed to describe the chaotic behavior of ray trajectories in a deep-water acoustic waveguide up to three to five thousands of kilometers in length. The ray dynamics is investigated using the Hamiltonian formalism expressed in terms of the canonical action–angle variables. A realistic waveguide model is used, with refractive-index fluctuations due to the random field of internal waves. The Fokker–Planck equation is obtained for the action variable, and it is shown that the range dependence of this variable can be approximated by the Wiener random process, which represents the simplest model of diffusion. Formulas are derived for calculating the probability density of the coordinate and other ray characteristics. An approximate expression is found for the smoothed field intensity of a point source. For illustrating and testing the formulas obtained, their predictions are compared with the results of numerical solutions of ray equations and the results of field calculations by the parabolic equation method. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The numerical modeling of the fields in underwater acoustic waveguides testifies that the ray theory provides correct predictions for many important characteristics of signals received even at ranges up to several thousands of kilometers long. Many results obtained in the geometrical optics approximation agree well with the computations performed by the parabolic equation method, as well as with the data of field experiments [1–4]. This stimulates interest in the development of new methods for analyzing the ray structure of the field under the conditions of long-range sound propagation in the ocean.

As is known, the horizontal scales of sound-speed variations in sea water considerably (by two orders of magnitude at minimum) exceed the vertical ones [5]. This fact allows one to ignore the horizontal ray refraction and, therefore, confine one's consideration to the analysis of the two-dimensional environmental model [6, 7]. In this case, the sound speed is given by the function $c(r, z)$, where r is the distance and z is the depth. In typical waveguide models, this function is represented in the form

$$c(r, z) = c_0(r, z) + \delta c(r, z), \tag{1}$$

where $c_0(r, z)$ is the unperturbed sound speed field smoothly depending on both coordinates and $\delta c(r, z)$ is a small perturbation. It is assumed that $\delta c(r, z)$ is a random function and, therefore, the study the ray structure of the sound field requires the application of statistical methods.

In the theory of wave propagation in random media, the description of rays (as well as other characteristics of wave fields) is based on the notion of a statistical ensemble of realizations of a medium. The statistical characteristics of a ray with initial parameters z_0 and χ_0 representing the initial depth and launch angle, respectively, are determined by averaging over the rays with the same initial conditions in various realizations forming the ensemble. This approach is the conventional one, and it is used in most studies concerned with the analysis of the statistical ray structure [8].

In the late 1980s, a new approach to studying the stochastic dynamics of rays was developed on the basis of the optical-mechanical analogy [4, 9–14]. The point is that the behavior of a ray trajectory in a range-dependent waveguide obeys almost the same equations that describe the oscillations of a nonlinear oscillator under the action of a nonstationary external force. It is well known that, in mechanics, the situation where such an oscillator exhibits a chaotic behavior is typical [15]. Numerical calculations confirmed that the ray trajectories in underwater waveguides behave, as expected, in a similar way [4, 16]. The chaotic rays are characterized by a high instability: for the trajectories with very close initial conditions, the difference in vertical coordinates Δz grows with distance r, on the average, according to the exponential law [9]

$$\Delta z \sim \exp(\lambda r). \tag{2}$$

The aim of this paper is the approximate analytical description of the statistical ray chaos in a realistic model of a deep-sea acoustic waveguide with sound-speed inhomogeneities $\delta c(r, z)$ caused by random internal waves. In this model, the Lyapunov exponents λ have values on the order of 0.01 km^{-1} [4]. At distances exceeding 1000 km, the ray chaos is fully developed

and the inclusion of this phenomenon in the analysis of the sound field statistics is necessary.

This paper investigates the chaotic ray dynamics in a determinate medium with spatial sound-speed fluctuations given by a single realization of the perturbation $\delta c(r, z)$. Using the fact that the ray trajectories with close initial conditions become almost independent at $r \gg \lambda^{-1}$, the averaging over initial conditions is be treated as a statistical one. The results of numerical modeling testify that the resulting statistical characteristics of chaotic rays weakly depend on the specific realization $c(r, z)$ used in the calculations. Therefore, one can expect that the analysis of ray statistics in a single realization of the sound-speed field should give the results close to those that could be obtained by averaging over a statistical ensemble of waveguides. However, the very important and interesting question of how well one realization of inhomogeneities represents the complete statistical ensemble is beyond the scope of this paper and is not considered here.

The method of the analysis of chaotic ray dynamics that is presented in this paper is based on the Hamiltonian formalism expressed in terms of the canonical action–angle variables [9, 17]. These variables are widely used for analyzing the finite motions of particles in mechanics, as well as for studying the oscillations of ray trajectories in waveguides. The action variable I determines the amplitude and the ray cycle length, while the angular variable θ (not to be confused with the geometric grazing angle of a ray) defines the position of the current point of a ray within the given cycle. One can say that θ plays the role of the phase of an oscillating trajectory.

For most rays, the variable θ soon becomes stochastic: even on rather short tracks, it can be considered as uniformly distributed over the interval from 0 to 2π. This allows one to obtain the Fokker–Planck equation for the action variable. A numerical calculation shows that, in our waveguide model, the diffusion coefficient I can be approximated by a constant. In this approximation, the random function $I(r)$ is modeled by the Wiener process and, therefore, the probability density of the action is determined by the well-known formula.

Knowing the combined probability density of I and θ, one can calculate almost any statistical ray characteristics. In this paper, the capabilities of this approach are illustrated by the example of the field produced by a point source. Two specific problems are considered. First, for chaotic rays launched from the point source, the probability densities of their parameter distributions are determined. Second, a formula is obtained for the field intensity in the vertical section of the waveguide with smoothing of the intensity by a smooth weighting function. The predictions obtained in the framework of our approximation (let us call it the Wiener process approximation) are compared with the results (numerical) of the direct calculations of the ray trajectories and the wave field.

In this study, the parabolic equation approximation (small-angle approximation) is used, although the generalization of the results to the case where the field is determined by the Helmholtz equation presents no serious problems. The point is that the parabolic equation, in contrast to the Helmholtz equation, can be easily solved numerically [5]. The formula for the smoothed field intensity is tested using one of the known codes for solving parabolic equations, which is called MMPE [18].

A more detailed description of the approach proposed here is presented in the electronic preprint (e-print) [19]. It can be found on the internet, on the site xxx.lanl.gov.

2. THE HAMILTONIAN FORMALISM FOR THE RAY TRAJECTORY DESCRIPTION

2.1. Momentum–Position Variables

For simplicity, we assume that the unperturbed sound-speed field in Eq. (1) does not depend on r; i.e., it is presented by a smooth function $c_0(z)$ along the whole acoustic path. Let the axis z be directed downward and the water surface coincide with the plane $z = 0$. The refractive index $n(r, z)$ equals $c_r/c(r, z)$, where c_r is an auxiliary constant, which in underwater acoustics can always be chosen so that $|c - c_r| \ll c_r$. In this case, $n(r, z) = n_0(z) + \delta n(r, z)$, where $n_0(z) = c_r/c_0(z)$ and $\delta n \simeq -\delta c/c_r$ is the small perturbation of the refractive index that is responsible for the ray chaos.

For the analysis of ray trajectories, we use the Hamiltonian formalism [9, 20]. In its framework, the ray equations have the form

$$\frac{dz}{dr} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dr} = -\frac{\partial H}{\partial z}, \tag{3}$$

where

$$H = \frac{p^2}{2} + \frac{(1 - n^2)}{2}. \tag{4}$$

The function H can be interpreted as the Hamiltonian of some (conceptual) mechanical system. The trajectory the parameter p involved in Eqs. (3) and (4) is an analog of the mechanical momentum, and the distance r plays the role of time. The relation between the momentum and the ray grazing angle χ is given by the formula $p = \tan\chi$. The expression for the Hamiltonian is presented in the small-angle approximation ($p \ll 1$), which, as stated in the introduction, is used in this paper.

2.2. Action–Angle Variables

In an underwater acoustic waveguide, the ray trajectories have the form of oscillating curves. For their description (as for the description of mechanical particle oscillations in a potential well), along with the momentum–position variables (p, z), it is convenient to

use the action–angle canonical variables (I, θ) [9, 17]. To determine the latter, we consider an unperturbed waveguide with the Hamiltonian

$$H_0 = \frac{p^2}{2} + \frac{(1 - n_0^2)}{2}. \quad (5)$$

According to Snell's law (an analog of the energy-conservation law in mechanics), the quantity H_0 remains constant along a ray path and, therefore, Eq. (5) defines the momentum p as the function of coordinate z:

$$p = \pm\sqrt{2H_0 - 1 + n_0^2}. \quad (6)$$

The action variable is given by the integral along the period of the unperturbed trajectory:

$$I = \frac{1}{2\pi}\oint p\,dz = \frac{1}{\pi}\int_{z_{\min}}^{z_{\max}} dz\sqrt{2H_0 - 1 + n_0^2(z)}. \quad (7)$$

Here, $z_{\min}$ and $z_{\max}$ are the depths of the lower and upper turning points of the ray. Integral (7) defines the "energy" H_0 as the function of the action variable I. An important characteristic of the waveguide is the derivative of this function

$$\frac{dH_0}{dI} = \frac{2\pi}{D} = \omega, \quad (8)$$

which determines the spatial frequency of trajectory oscillations. In this formula, D is the ray cycle length (period).

The known generating function of the canonical transformation that relates the pairs of variables (p, z) and (I, θ) [9, 17] can be presented in the form

$$G(I, z) = \begin{cases} \displaystyle\int_{z_{\min}}^{z} dz\sqrt{2H_0(I) - 1 + n_0^2(z)}, & p > 0 \\ \displaystyle 2\pi I - \int_{z_{\min}}^{z} dz\sqrt{2H_0(I) - 1 + n_0^2(z)}, & p < 0, \end{cases} \quad (9)$$

where $z_{\min}$ and $z_{\max}$ are considered to be functions of I. The canonical transformation given by the functions

$$z = z(I, \theta), \quad p = p(I, \theta) \quad (10)$$

(both functions are periodic in θ with a period of 2π) and

$$I = I(p, z), \quad \theta = \theta(I, z) \quad (11)$$

is determined by the equations

$$p = \frac{\partial G}{\partial z}, \quad \theta = \frac{\partial G}{\partial I}. \quad (12)$$

By virtue of the Liouville theorem [17], the Jacobian is

$$\frac{\partial(z, p)}{\partial(I, \theta)} = 1. \quad (13)$$

In spite of the fact that this transformation is determined for an unperturbed range-independent waveguide, it can formally be used in the presence of perturbations as well. Substituting Eqs. (10) into Eq. (4), we represent the perturbed Hamiltonian as

$$H(I, \theta, r) = H_0(I) + V(I, \theta, r), \quad (14)$$

where

$$V \simeq -\delta n \simeq \delta c/c_r. \quad (15)$$

In terms of the action–angle variables, the Hamiltonian equations assume the form

$$\frac{dI}{dr} = -\frac{\partial V}{\partial \theta}, \quad (16)$$

and

$$\frac{d\theta}{dr} = \omega + \frac{\partial V}{\partial I}. \quad (17)$$

In closing this section, it should be noted that the canonical transformation given by generating function (9) is expressed in terms of elementary or special functions only for several simplest profiles $n_0(z)$. However, this transformation can easily be realized numerically with a standard code for calculating the ray trajectories in a range-independent waveguide. This issue is discussed in [19].

3. STATISTICAL CHARACTERISTICS OF CHAOTIC RAYS

3.1. The Fokker–Planck Equation

Let us consider Hamiltonian equations (16) and (17) as the Langevin stochastic equations. In the case of weak inhomogeneities, the typical horizontal scale l_m of the function $\delta c(r, z)$ is much smaller than the scale l_I of the action variable variations:

$$l_m \ll l_I. \quad (18)$$

In this case, the random function on the right-hand side of Eq. (16) can approximately be considered as a delta-correlated function. Then, $I(r)$ is a Markovian process whose probability density $P(I, r)$ obeys the Fokker–Planck equation [21, 22]

$$\frac{\partial P}{\partial r} = -\frac{\partial}{\partial I}(AP) + \frac{1}{2}\frac{\partial^2}{\partial I^2}(BP). \quad (19)$$

The coefficients of this equation, A and B, are determined by the relationships

$$A(I, r) = \lim_{\Delta r \to 0}\frac{1}{\Delta r}\int \Delta I P(I + \Delta I, r + \Delta r | I, r)\,d\Delta I, \quad (20)$$

$$B(I, r) = \lim_{\Delta r \to 0}\frac{1}{\Delta r}\int \Delta I^2 P(I + \Delta I, r + \Delta r | I, r)\,d\Delta I, \quad (21)$$

where $P(I, r|I_0, r_0)$ is the probability density of the action I at the distance r, provided that the action magnitude is I_0 at the point $r_0 \le r$.

By virtue of the detailed balance principle, which expresses the symmetry of the Hamiltonian equations with respect to the time reversal (in our case, a change of the sign of r), the following relationship is valid [22]:

$$A = \frac{1}{2}\frac{\partial B}{\partial I}. \quad (22)$$

In this case, the Fokker–Planck equation takes the form

$$\frac{\partial P}{\partial r} = \frac{1}{2}\frac{\partial}{\partial I}B\frac{\partial}{\partial I}P. \quad (23)$$

The derivation of this equation without using the detailed balance principle is given in [19]. To calculate the coefficient B, we choose two points r and $r_0 < r$ in such a way that

$$l_m \ll r - r_0 \ll l_I. \quad (24)$$

We denote the action and angle variables at the point r_0 as I_0 and θ_0, respectively, and their values at the point r, as I and θ. According to condition (24), we can use the approximate equality

$$\Delta I = I - I_0 = -\int_{r_0}^{r} dr' V_\theta(I_0, \theta^{(0)}(r'), r'), \quad (25)$$

where

$$\theta^{(0)}(r) = \theta_0 + \omega(I_0)(r - r_0) \quad (26)$$

is the unperturbed solution to Eq. (17). Let us introduce the correlation function

$$K(I, \rho) = \langle V_\theta(I, \theta^{(0)}(r+\rho), r+\rho) V_\theta(I, \theta^{(0)}(r), r)\rangle_{\theta_0, V}, \quad (27)$$

where the symbol $\langle \ldots \rangle_{\theta_0, V}$ means averaging over θ_0 and over inhomogeneities V that are located between r_0 and r. Assume that the medium is statistically homogeneous along the r axis and, therefore, the function K does not depend on r. In this case, we have

$$\langle \Delta I^2 \rangle_{\theta_0, V} = (r - r_0)\int_0^{r - r_0} d\rho K(I, \rho). \quad (28)$$

Condition (24) allows one to extend the upper limit to infinity. On the other hand, inequality (24) means that $r - r_0$ is small compared to the scale of variation of I. Therefore, it is possible to substitute Eq. (28) into Eq. (21) by replacing $(r - r_0)$ with Δr. Using the property $K(I, \rho) = K(I, -\rho)$, we obtain

$$B(I) = \frac{1}{2}\int_{-\infty}^{\infty} d\rho K(I, \rho). \quad (29)$$

If the angular variable θ becomes stochastic (i.e., uniformly distributed over the interval from 0 to 2π) at distances that are small compared to the distance to the observation point, the function $B(I)$ can be found numerically from a single realization of the inhomogeneous medium. From Eq. (15), it follows that

$$V_\theta = -\frac{\partial \delta n}{\partial z}\frac{\partial z}{\partial \theta}, \quad (30)$$

where the derivative with respect to θ is calculated for the function $z(I, \theta)$—one of the functions determining canonical transformation (10). In a range-independent waveguide, for small values of p, we have

$$\frac{\partial z}{\partial \theta} = \frac{1}{\omega(I)}\frac{dz}{dr} = \frac{1}{\omega(I)}\frac{p}{\sqrt{n^2 - p^2}} \simeq \frac{p}{\omega(I)}. \quad (31)$$

When operating with one realization of inhomogeneities $\delta n(r, z) = -\delta c(r, z)/c_r$, for the numerical calculation of correlation function (27) we consider an unperturbed ray trajectory with the action variable I. Let the coordinate and the momentum along this trajectory be given by functions $z(r)$ and $p(r)$, respectively. According to Eqs. (30) and (31), at a distance r, the derivative V_θ at a point on the trajectory of the unperturbed ray is equal to

$$V_\theta = -\delta n_z(r)\frac{p(r)}{\omega}, \quad (32)$$

where

$$\delta n_z(r) = \left.\frac{\partial \delta n(r, z)}{\partial z}\right|_{z = z(r)}. \quad (33)$$

The estimate of the correlation function is given by the integral along the trajectory of the unperturbed ray

$$K(I, \rho) = \frac{1}{R\omega^2(I)}\int_0^R dr \delta n_z(r + \rho) p(r + \rho) \delta n_z(r) p(r) \quad (34)$$

on a track whose length R must be much greater than both the ray cycle length and the maximal horizontal scale of inhomogeneities. A more accurate estimate of the correlation function can be obtained by calculating integral (34) along a group of unperturbed rays with the same values of the action variable I but with different initial values of the angular variable θ and with a subsequent averaging of the results obtained. In the next section, a specific example of calculating the diffusion coefficient $B(I)$ by the aforementioned method is presented.

3.2. The Action Variable in the Wiener Process Approximation

Consider a model of a deep-water acoustic waveguide with an unperturbed sound-speed profile $c_0(z)$ shown at the left of Fig. 1. Weak fluctuations of the sound speed $\delta c(r, z)$ are caused by the effect of the random field of internal waves, whose statistics is determined by the Garrett–Munk empirical spectrum [6]. In every point of the waveguide, the average (over the

ensemble of inhomogeneities) value of perturbations is zero. For the numerical modeling of specific field realizations $\delta c(r, z)$, we use the method proposed in [23]. Omitting details, note that the rms amplitude of $\delta c(r, z)$ near the sea surface equals 0.5 m/s and decreases with depth by the exponential law $\exp(-z/L)$, where L = 0.66 km. As a function of the horizontal coordinate r, the spectrum of perturbations is concentrated in the interval of spatial frequencies from $2\pi/100$ km^{-1} to $2\pi/4$ km^{-1}. The right-hand side of Fig. 1 exhibits the vertical sections of the field $\delta c(r, z)$ at three different distances. In such a waveguide model, the rays behave chaotically [19, 24].

The dependence of the diffusion coefficient B on the action I, as calculated from formulas of Section 3.1, is shown in Fig. 2. This dependence proves to be rather weak, and the quantity B can be approximated by a constant:

$$B = 1.5 \times 10^{-7} \text{ km}. \quad (35)$$

In what follows, this approximation will be called the Wiener process approximation. The point is that, at a constant B, the variable I can be represented as

$$I = I_s + x, \quad (36)$$

where $I_s = I(0)$ is the initial value of the action and $x(r)$ is the Wiener random process that begins at the point $x(0) = 0$ [21, 22]. The process $x(r)$ obeys the stochastic equation

$$\frac{dx}{dr} = \xi(r), \quad (37)$$

where ξ is the white noise whose mean value and correlation function are given by the formulas

$$\langle \xi \rangle = 0, \quad \langle \xi(r)\xi(r') \rangle = B\delta(r - r'). \quad (38)$$

Consider the following important point. The random function $I(r)$ given by relationships (36)–(38) can assume both positive and negative values. However, the action variable is nonnegative by definition. This fact can easily be taken into account by introducing a reflecting boundary at $I = 0$. The conditional probability density $P_{I|I_0}(I, r|I_s)$ at the distance r, i.e., the probability density of the variable I on condition that its initial value (at $r = 0$) is equal to I_s, is given by the solution to Fokker–Planck equation (23) with a constant diffusion coefficient and with the initial condition $P(I, 0|I_s) = \delta(I - I_s)$. In the presence of the reflecting boundary, the desired solution is [21]

$$P_{I|I_0}(I, r|I_s) = \frac{1}{\sqrt{2\pi Br}}\left[\exp\left(-\frac{(I-I_s)^2}{2Br}\right) + \exp\left(-\frac{(I+I_s)^2}{2Br}\right)\right]. \quad (39)$$

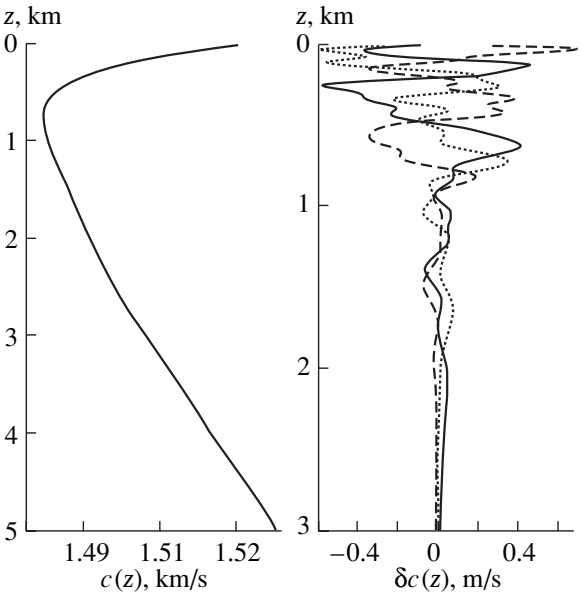


Fig. 1. Unperturbed sound-speed profile $c_0(z)$ (at the left) and vertical sections of the perturbation $\delta c(r, z)$ for three different distances (at the right).

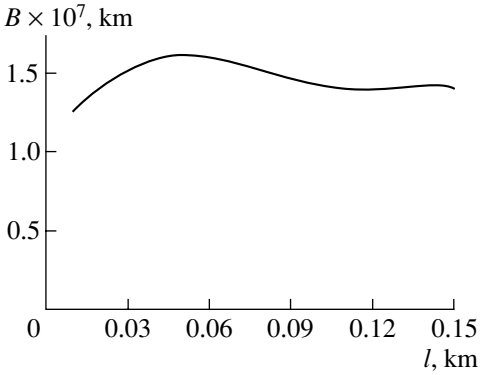


Fig. 2. Dependence of the diffusion coefficient B on the action I.

Our method of describing the chaotic ray statistics is based on this formula.

The magnitude of the action variable I_s grows as the launch angle χ_s of a ray from the source increases (this is the grazing angle of the ray, which should not be confused with the canonical angular variable θ). For steep rays with

$$I_s \gg \sqrt{Br}, \quad (40)$$

the second term in the square brackets in Eq. (39) can be neglected. This approximation means that the trajectories of the Wiener process on the track under consideration do not reach the reflecting boundary previously introduced by us. In this case, the description of the ray

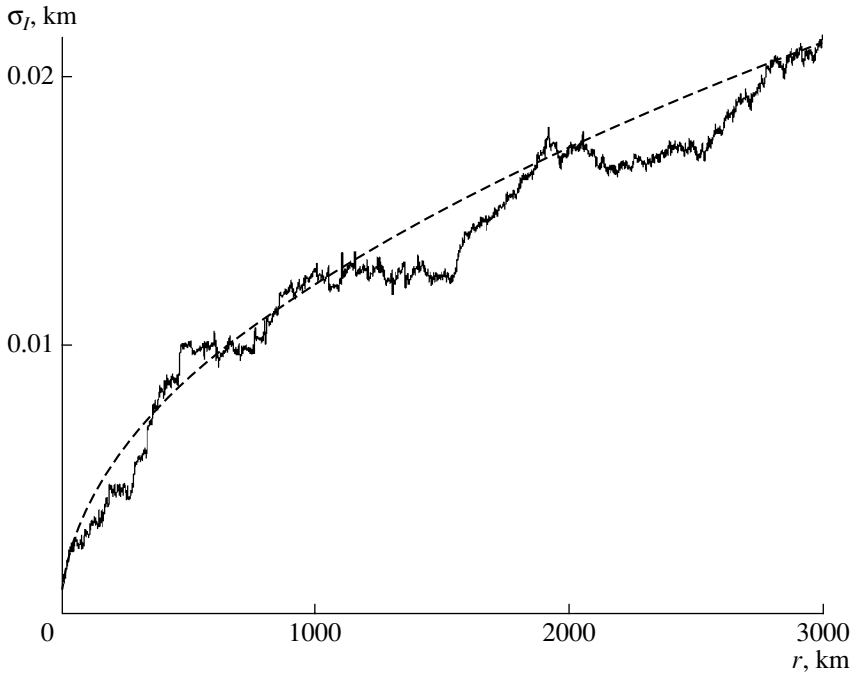


Fig. 3. Standard deviation of the action variable I as a function of distance r: the result of averaging over 100 ray trajectories with the initial action $I_S = 0.06$ km and different initial values of the angle variable θ (the solid curve); the Wiener process approximation (the dashed curve).

statistics is considerably simplified. In particular, the quantity I spreads with distance according to the diffusion law

$$\sigma_I = \sqrt{\langle (I - I_s)^2 \rangle} = \sqrt{Br}. \quad (41)$$

The results of the numerical modeling show that criterion (40) can be moderated by replacing the symbol ≫ with >. For a point source located at the waveguide axis, this condition is met at the distance $r = 3000$ km for rays with the launch angles $\chi_s > 5°$. Consider a numerical example. Figure 3 shows the standard deviation of the variable I as a function of distance for a ray with the initial value of the action variable $I_s = 0.06$ km. In an unperturbed waveguide, such a ray crosses the waveguide axis under the grazing angle $\chi_s = 7.8°$. The solid line shows the result of averaging over 100 rays with the aforementioned initial value of the action variable and initial values of the angular variable, which are uniformly distributed over the interval from 0 to 2π. The dashed curve is given by Eq. (41). As is seen, the simplest statistical model constructed in this section for describing the fluctuations of I agrees well with the results of the direct numerical modeling. Similar results are obtained for rays with other initial conditions.

Equation (37) gives a simplified form of Hamilton equation (16). The simplified form of the second Hamilton equation (17) is considered in detail in [19].

4. THE FIELD OF A POINT SOURCE

4.1. Probability Densities of Chaotic Ray Parameters

The probability density $P_{I|I_0}(I, r|I_s)$ describes the distribution of the action variable for an ensemble of rays with initial values of the angular variable θ_s uniformly distributed over the interval from 0 to 2π. Now, we consider a point source located at the point $(0, z_s)$. We assume that the initial momenta of rays, p_0, are within the interval

$$-p_{\max} < p_0 < p_{\max}. \quad (42)$$

We consider the chaotic rays with such initial conditions as a statistical ensemble, the description of which is our aim. In this case, the probability for an arbitrary ray to have the initial momentum in a small interval $(p_0, p_0 + dp_0)$ is $dp_0/(2p_{\max})$. For the rays with initial momenta from this interval, the probability density of the action at a distance r is $P_{I|I_0}(I, r|I(p_0, z_s))$ (here, we use the function $I(p, z)$ that appears on the right-hand side of the first of Eqs. (11)). The probability density of the action for the rays with momenta lying within interval (42) is given by the integral

$$P_I(I, r) = \frac{1}{2p_{\max}} \int_{-p_{\max}}^{p_{\max}} dp_0 P_{I|I_0}(I, r|I(p_0, z_s)). \quad (43)$$

In the Wiener process approximation, the angular variable θ is assumed to be uniformly distributed over the

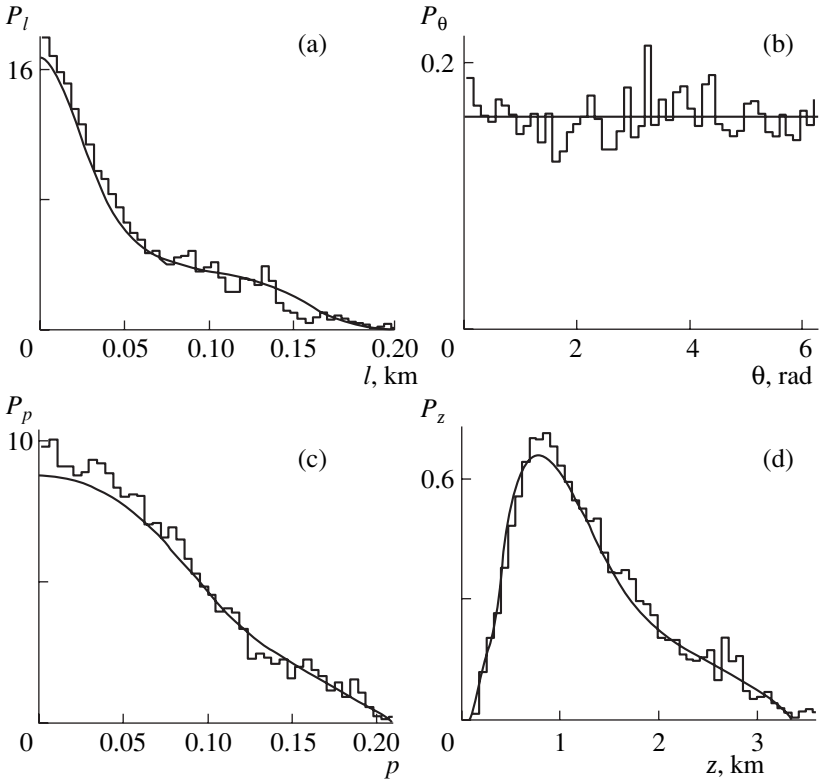


Fig. 4. Probability densities (distributions) of the variables (a) I, (b) θ, (c) p, and (d) z at a distance of 3000 km for rays launched from a point source located at a depth of 0.78 km—the Wiener process approximation (the solid curves). The stepped lines represent the normalized histograms of the distributions obtained by a numerical calculation over 48000 trajectories.

interval $(0, 2\pi)$ and statistically independent of I. Then, the combined probability density of I and θ is

$$P_{I\theta}(I, \theta, r) = \frac{1}{2\pi} P_I(I, r). \quad (44)$$

Using standard formulas of probability theory and relationship (13), from Eq. (44) we obtain the formula for the joint probability density of p and z:

$$P_{pz}(p, z, r) = \frac{1}{2\pi} P_I(I(p, z), r). \quad (45)$$

From Eq. (45), we find the probability densities of both momentum

$$P_p(p, r) = \frac{1}{4\pi p_{\max}} \int_{-p_{\max}}^{p_{\max}} dp_0 dz P_{I|I_0}(I(p, z), r | I(p_0, z_s)) \quad (46)$$

and the coordinate

$$P_z(z, r) = \frac{1}{4\pi p_{\max}} \int_{-p_{\max}}^{p_{\max}} dp_0 dp P_{I|I_0}(I(p, z), r | I(p_0, z_s)). \quad (47)$$

The smooth curves in Fig. 4 show the probability densities for I, θ, p, and z on a 3000-km-long track, which are calculated in the Wiener process approximation, i.e., by formulas (43), (46), and (47) (in this approximation, θ is expected to be uniformly distributed). These curves are compared with the estimates of the probability densities obtained by the numerical calculation for 48000 rays. The calculations are carried out for a point source positioned at a depth of 0.78 km. The boundaries of interval (42) were determined from the condition that the initial grazing angles of rays are within the range ±12°. The stepped lines in Fig. 4 show normalized histograms of the distributions of I, θ, p, and z along the given track. As is seen, the predictions based on the Wiener process approximation agree well with the results of the numerical modeling.

4.2. Smoothed Distribution of the Field Intensity

In this paper, we assume that the complex amplitude of the field is given by the parabolic equation [5]

$$2ik\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} - k^2(1 - n^2(r, z))u = 0, \quad (48)$$

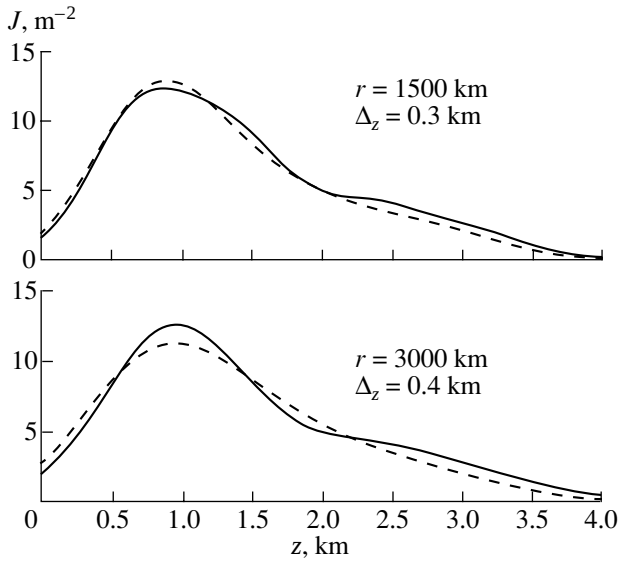


Fig. 5. Smoothed sound field intensity of a point source in the vertical section of the waveguide at distances of 1500 (the upper plot) and 3000 km (the lower plot): the numerical calculation with the parabolic equation method (the solid curves) and the Wiener process approximation (the dashed curves). The source depth is 0.78 km, and the carrier frequency is 75 Hz. The smoothing scales Δ_z are indicated in the plots.

where $k = 2\pi f/c_r$ and f is the carrier frequency. The field of the point source considered in the previous section obeys equation (48) with the initial condition

$$u(0, z) = \delta(z - z_s). \tag{49}$$

Equation (48) coincides with the Schrödinger equation of quantum mechanics, but, in our case, the role of the Planck constant is played by k^{-1}. In the geometrical optics approximation, $u(r, z)$ is given by the superposition of ray arrivals at the point of observation.

When analyzing the sound fields in acoustic waveguides, the common practice is to limit the calculations to a rough estimate of the field intensity by an incoherent summation of rays [5]. This is explained not only by the mere simplification of calculations. The calculation of fine details of the wave pattern (the knowledge of which is often unnecessary in practice) under the conditions of multipath propagation typical of waveguide problems is complicated by inevitable inaccuracies of the mathematical model of the medium and the approximate character of geometrical optics. The incoherent summation of rays gives an estimate averaged over spatial coordinates for the field intensity distribution in the waveguide. In this case, the dependence of the result on small variations of the waveguide parameters is weaker than in the case of the calculation with allowance for the ray phases. In the problem under study, the incoherent summation of rays leads to the expression

$$|u(r, z)|^2 = \sum_m \frac{k}{2\pi|\partial z/\partial p_0|}\bigg|_{p_0 = p_{0,m}}, \tag{50}$$

where $p_{0,m}$ is the initial momentum of the mth ray arriving at the point (r, z). In Eq. (50), the known formula for the ray amplitude in the parabolic equation approximation is used [25].

Under the ray chaos conditions, the number of rays arriving at the observation point becomes very large. Therefore, the use of formula (50) requires many calculations. In addition, the result will widely differ from the true field intensity at the point (r, z). The situation is radically simplified if the incoherent summation of rays is complemented with spatial averaging. Here, we consider the smoothed intensity defined by the relationship

$$J(r, z) = \frac{1}{\sqrt{2\pi}\Delta_z}\int dz' |u^2(r, z')| \exp\left(-\frac{(z - z')^2}{2\Delta_z^2}\right). \tag{51}$$

In [26, 27], it was conjectured that the choice of sufficiently large scales of smoothing Δ_z removes the difference between the exact value of J obtained from the parabolic equation solution and the approximate value of this quantity found by the incoherent summation of rays, i.e., from the combination of Eqs. (50) and (51). This assumption was confirmed by the results of numerical modeling. Unfortunately, the question about the choice of necessary scales of smoothing is as yet poorly investigated.

In the Wiener process approximation, one can obtain the analytical description of the smoothed field intensity. Substituting Eq. (50) into Eq. (51), we obtain $J(r, z)$ in the form

$$J(r, z) = \frac{k}{(2\pi)^{3/2}\Delta_z}\int_{-p_{\max}}^{p_{\max}} dp_0 \exp\left(-\frac{(z - z(r, p_0, z_s))^2}{2\Delta_z^2}\right), \tag{52}$$

where the function $z(r, p_0, z_s)$ denotes the trajectory of the ray launched from the point z_s with the initial momentum p_0. Under the ray chaos conditions, the integration over the initial conditions is equivalent to a statistical averaging. Therefore, using Eq. (47), we approximately replace Eq. (52) by

$$J(r, z) = \frac{kp_{\max}}{(2\pi)^{3/2}\Delta_z}\int_{-p_{\max}}^{p_{\max}} dz' \exp\left(-\frac{(z - z')^2}{2\Delta_z^2}\right) P_z(z', r). \tag{53}$$

Thus, we obtained the analytical expression for the smoothed field amplitude under the ray chaos conditions. Figure 5 indicates that formula (53) adequately predicts the results of the numerical calculation of $J(r, z)$ by the direct solution of the parabolic equation. The

wave field $u(r, z)$ was computed with the MMPE code [18]. The scales of smoothing Δ_z that were used to construct the plots were chosen empirically.

5. CONCLUSIONS

This paper discusses an analytical approach to the approximate description of the chaotic ray dynamics in a deep-water acoustic waveguide with inhomogeneities induced by the random field of internal waves. The method is based on using the Hamiltonian formalism in terms of the action–angle canonical variables (I, θ). The analysis of ray statistics is considerably simplified by the fact that the angular variable θ can be considered as uniformly distributed over the interval $(0, 2\pi)$ even at comparatively short ranges. In this approximation, the dependence of the action I on distance r is well approximated by the Wiener random process, which represents the simplest diffusion model [21].

Canonical transformation (10) from (I, θ) to (p, z) can be interpreted as a nonlinear change of variables. It allows one to recalculate the known statistical characteristics (I, θ) as the characteristics (p, z) by using standard formulas of probability theory. According to the Liouville theorem, the Jacobian of the transformation from (I, θ) to (p, z) equals unity, which simplifies the recalculation process. As a result, expressions are obtained for analyzing the statistics of the coordinates and grazing angles of chaotic rays.

It should be noted that, in the Wiener process approximation, the influence of the environmental inhomogeneities is determined by only one parameter: the action diffusion coefficient B. The canonical transformation used for the recalculation of statistical characteristics is determined by only the unperturbed sound-speed profile.

As an example of the application of the results obtained, the probability densities of the variables I, p, and z are calculated for the rays launched from a point source. In addition, an analytical expression is obtained for the smoothed acoustic intensity. The results of numerical modeling agree well with the estimates obtained with our approximate formulas.

All calculations described in this paper, including the calculation of the diffusion coefficient of the action variable, are carried out for a single realization of inhomogeneities. Here, an analog of statistical averaging is the averaging over the initial conditions. Numerical calculations testify that the ray statistics weakly varies with a change from one realization to another. However, the question of the extent of general nature of the results of studying the ray dynamics for a single realization of inhomogeneities remains open and requires further investigation.

The approach considered above can be used to study the statistics of other parameters of chaotic rays. In particular, this refers to the travel times of sound pulses along ray trajectories connecting the source and the receiver (for brevity, these times are called travel times). In [19, 24], it is shown that the variations of ray arrival times are easily expressed in terms of the variations of the action variable. Therefore, our formulas for calculating the statistical characteristics of I allow one to describe the arrival-time fluctuations [19]. In particular, in [19], a quantitative theory is developed to describe the effect of separation of the travel times of chaotic rays into compact clusters and an explanation is offered for the unexpected high stability of the initial part of a sound pulse in a deep sea, which is observed up to distances of several thousands of kilometers.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no.03-02-17246) and by a Scientific School grant (no. NSh-838–2003).

REFERENCES

1. P. F. Worcester, B. D. Cornuelle, M. A. Dzieciuch, *et al.*, J. Acoust. Soc. Am. **105**, 3185 (1999).
2. J. A. Colosi, E. K. Scheer, S. M. Flatte, *et al.*, J. Acoust. Soc. Am. **105**, 3202 (1999).
3. B. D. Dushaw, B. M. Howe, J. A. Mercer, and R. C. Spindel, IEEE J. Ocean. Eng. **24**, 202 (1999).
4. J. Simmen, S. M. Flatte, and G.-Y. Wang, J. Acoust. Soc. Am. **102**, 239 (1997).
5. L. M. Brekhovskikh and Yu. P. Lysanov, *Fundamentals of Ocean Acoustics* (Gidrometeoizdat, Leningrad, 1982; Springer, New York, 1991).
6. *Sound Transmission through a Fluctuating Ocean*, Ed. by S. M. Flatte (Cambridge Univ. Press, Cambridge, 1979; Mir, Moscow, 1982).
7. *Wave Propagation and Underwater Acoustics*, Ed. by J. B. Keller and J. S. Papadakis (Springer, New York, 1977; Mir, Moscow, 1980).
8. S. M. Rytov, Yu. A. Kravtsov, and V. I. Tatarskiĭ, *Introduction to Statistical Radiophysics*, Part 2: *Random Fields* (Nauka, Moscow, 1978) [in Russian].
9. G. M. Zaslavskiĭ and S. S. Abdullaev, Usp. Fiz. Nauk **161** (8), 1 (1991) [Sov. Phys. Usp. **34**, 645 (1991)].
10. D. R. Palmer, M. G. Brown, F. D. Tappert, and H. F. Bezdek, Geophys. Res. Lett. **15**, 569 (1988).
11. K. B. Smith, M. G. Brown, and F. D. Tapper, J. Acoust. Soc. Am. **91**, 1939 (1992).
12. K. B. Smith, M. G. Brown, and F. D. Tappert, J. Acoust. Soc. Am. **91**, 1950 (1992).
13. I. P. Smirnov, A. L. Virovlyansky, and G. M. Zaslavsky, Phys. Rev. E **64**, 036221 (2001).
14. M. G. Brown, J. A. Colosi, S. Tomsovic, *et al.*, J. Acoust. Soc. Am. **113**, 2533 (2003).
15. G. M. Zaslavskiĭ and R. Z. Sagdeev, *Introduction to Nonlinear Physics: From a Pendulum to Turbulence and Chaos* (Nauka, Moscow, 1988) [in Russian].

16. M. G. Brown and J. Viechnicki, J. Acoust. Soc. Am. **104**, 2090 (1998).

17. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics*, 3rd ed. (Nauka, Moscow, 1973; Pergamon Press, Oxford, 1976).

18. K. B. Smith, J. Comput. Acoust. **9**, 243 (2001).

19. A. L. Virovlyansky, *Ray Travel Times in Range-Dependent Acoustic Waveguides*, nlin. CD/0012015.

20. M. Born and E. Wolf, *Principles of Optics*, 4th ed. (Pergamon Press, Oxford, 1969; Nauka, Moscow, 1973).

21. V. I. Tikhonov and M. A. Mironov, *The Markovian Processes* (Sovetskoe Radio, Moscow, 1977) [in Russian].

22. C. Gardiner, *Handbook of Stochastic Methods in Physics, Chemistry, and Natural Sciences* (Springer, Heidelberg, 1983; Mir, Moscow, 1986).

23. J. A. Colosi and M. G. Brown, J. Acoust. Soc. Am. **103**, 2232 (1998).

24. A. L. Virovlyansky, J. Acoust. Soc. Am. **113**, 2523 (2003).

25. M. C. Gutzwiller, J. Math. Phys. **8**, 1979 (1967).

26. A. L. Virovlyansky and G. M. Zaslavsky, Chaos **10**, 211 (2000).

27. A. L. Virovlyanskiĭ and I. A. Okomel'kova, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **40** (12), 1542 (1997).

Translated by Yu. Lysanov

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 81–88. From Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 101–110.

Acoustoelectronics: History, Present State, and New Ideas for a New Era[1]

Yu. V. Gulyaev* and F. S. Hickernell**

**Institute of Radioengineering and Electronics, Russian Academy of Sciences, Moscow, 101999 Russia*
e-mail: gulyaev@cplire.ru
***University of Arizona, USA*
Received September 9, 2004

Abstract—The application of high-frequency acoustic devices to the enhancement of electronics saw an extraordinary growth in both Eastern and Western countries in the sixties and seventies. A major impetus for these developments was the tension existing between the Soviet Bloc countries in the east and the former Allied countries in the west. Government military spending on both sides provided funding to explore new acoustoelectronic concepts in universities, institutes, and major defense companies. The direct exchange of visits between scientists and engineers of the East and West was limited until the 1980s, when travel restrictions were lifted on both sides and authors that has previously only been names in the open literature became face-to-face contacts and enjoyed exchanges at conferences of mutual interest. This resulted in a new era of cooperative work between the East and West and a large number of device applications that are seen in electronic systems around the world today. This paper explores the major acoustoelectronic developments of the sixties and seventies from an eastern and western perspective. © *2005 Pleiades Publishing, Inc.*

When we accepted the invitation to make a historical report, we were aware that, at full scale, this undertaking is very difficult. It is difficult not only because acoustoelectronics, especially its surface acoustic waves part, is a rather wide branch of current science and technology. The main difficulty is that many important results were obtained in parallel in several countries and even in different firms and universities, and thus, in many cases, who was the first remains a big question. This especially relates to work carried out under conditions of secrecy in counterpoised Western and Eastern blocs of countries during the Cold War. Of course, there is also the danger of forgetting to mention someone, which may be painful.

Thus, we decided the following:

First: to restrict ourselves only to the time period from the beginning of the 1970s.

Second: to speak mainly on the physical foundations of acoustoelectronics rather than on its practical applications.

Third: to ask you to excuse us for the inevitable personal character of our estimates: we see events from the perspective of our age, which is over 70….

Acoustic waves of high frequency (>20 kHz, ultrasound) have been widely used for a long time in various branches of science and technology. Two important features of these acoustic waves are their relatively low propagation velocity (100000 times smaller than the velocity of light) and the ease of their excitation in piezoelectric materials with high efficiencies. This provided for their use in radio engineering and electronics. The delay lines with bulk acoustic waves have been used in radioengineering for many decades. Equally well known is another device using bulk acoustic waves in piezoelectric materials, namely, a quartz resonator for frequency stabilization. These two devices are very well-known examples of the use of acoustic waves (ultrasound) in radioelectronic systems for the processing and transmission of information signals.

Many scientists and engineers have contributed to this field—see the well-known books by W. Cady [1], W.P. Mason [2], B. Auld [3], and others. Since they are not very well known in the West, we would especially like to stress the pioneering contributions of the following Soviet (Russian) scientists: I.G. Shaposhnikov, who in 1941 first considered the propagation of bulk acoustic waves in piezoelectric materials [4]; A.V. Shubnikov, who investigated the piezoelectric properties of quartz and made the first quartz resonators [5]; and S.Ya. Sokolov, A.G. Sokolinsky, S.G. Kalashnikov, V.A. Krasil'nikov, G.K. Ul'yanov, S.S. Karinsky, K.N. Kozlovsky, L.K. Zarembo, L.D. Rosenberg, A.G. Smagin, M.I. Yaroslavsky, V.S. Bondarenko, I.A. Victorov, V.E. Lyamov, and many others.

The effects of the interaction of acoustic waves with free electrons, "acoustoelectronic effects," were probably first studied by I.G. Shaposhnikov in 1941 [4] for piezoelectric materials. He investigated the "electronic" absorption of acoustic waves and the change of the acoustic wave velocity due to an interaction with electron plasma.

[1] This article was submitted by the authors in English.

In 1953, Parmenter [6] predicted a third acoustoelectronic effect, the so-called acoustoelectric (AE) effect, consisting of the appearance of a dc voltage or a dc electric current in the direction of the acoustic wave propagation. Physically, the AE effect is due to the transmission by the acoustic wave of some part of its momentum to the electron gas in the process of electronic absorption. This results in the appearance of a dc electric current (acoustoelectric current) or a dc electric voltage (acoustoelectric voltage) if the specimen is open-circuited. Later, these effects were studied by H.E. Bommel [7], A.B. Pippard [8], G. Weinreich [9], and many others (see also the review paper by I.M. Lifshitz and M.I. Kaganov [10]).

A key point was the experimental observation by A.R. Hutson, J.H. McFee, and D.L. White in 1961 of the amplification of bulk ultrasonic waves in piezoelectric semiconductor CdS by the supersonic drift of electrons [11]. The very idea of the possibility of such an amplification was earlier discussed (in 1956) by the Russian scientists K.B. Tolpygo and Z.I. Uritsky and the American scientist G. Weinreich [12, 13]. The detailed theory of acoustic-wave (AW) amplification in semiconductors of various types, namely, piezoelectric, nonpiezoelectric, many-valley, with hot electrons, magnetic, etc., was developed in 1962 [14–22] (see also the review papers [23–25]).

Soon after this work [11], R.W. Smith [26] observed one more acoustoelectronic effect that is, in fact, a combination of previously described effects: the sublinearity and even the saturation of the current–voltage curve of the specimen under the conditions of acoustic-wave amplification. Physically, this effect is connected with the loss of the directed momentum of drifting electrons to acoustic fluctuations (phonons) amplified by them, which leads to a reduction of the drift current.

The very beautiful effect of acoustic-wave amplification brought to this field many brilliant scientists and engineers, and the number of publications grew rapidly. Everybody was thinking of the creation of a "solid-state traveling-wave tube" and other new microwave devices.

However, the researchers very soon understood that, since the mobility of electrons in good piezoelectric semiconductors of the CdS type is rather small (200 $cm^2/V\ s$), for good amplification it is necessary to apply a high voltage (several kilovolt). As a consequence, it was possible to use only a pulse operation regime, not dc. Otherwise, there is either a strong buildup of noise or the crystal burns out.

The solution to the problem was found in 1964 by Yu.V. Gulyaev and V.I. Pustovoit [27], who suggested the use of surface acoustic waves (SAWs) for amplification and proposed a layered-structure semiconductor–piezoelectric as the basic structure for an acoustic-wave amplifier and other possible acoustoelectronic devices with SAWs. This was indeed another key point. The SAW propagates along the surface of a solid and, thus, is accessible all along the path of its propagation. If the solid is piezoelectric, the SAW is accompanied by an electric field wave (in the general case, both longitudinal and transverse, perpendicular to the surface). Due to this field, the wave can interact with electrons in conducting media that are adjacent to the surface, excite electric currents, create a certain distribution of potentials, produce bunching of electrons, etc. On the other hand, the redistribution of potentials in the conducting media can have an influence on the SAW via the inverse piezoeffect. In particular, if there is a supersonic drift of electrons in the conducting media in the direction of SAW propagation, it may lead to the amplification of the SAW just as in a traveling-wave tube. Since there now appears the possibility of a wide choice of materials, one can take for this "layered structure" a very strong piezoelectric, not bothering about its conductivity, and a semiconductor with very high electron mobility, such as, for example, Ge, Si, or InSb. Thus, the dc regime of the SAW amplifier becomes possible.

In fact, the main idea of [27] is the indication that a SAW in a piezoelectric dielectric can interact with the electric fields and electrons from another conducting medium (one that is adjacent to the surface) all along the path of its propagation.

The further development of this idea was done in the work of R.M. White and F. Voltmer in 1965 [28], in which they suggested exciting the SAW electrically by a periodic interdigital structure of metallic electrodes on the surface of the piezoelectric with a period equal to the SAW wavelength, to which the alternating voltage is applied with the period of the wave. This construction, which is an analog of the Udo–Yagi antenna in electrodynamics and which was called an "interdigital transducer" in the simplest bidirectional excitation case, transforms the electric signal to a SAW very effectively (if it is infinitely long) with a loss of about 3 dB. Specially constructed unidirectional transducers can have losses of around 1 db or less. This method of SAW excitation is much more effective than previous attempts to mechanically excite SAWs by bulk acoustic waves with the help of wedges, grooves on the surface, combs, etc. (see, for example, [29]).[2]

Experimentally, the amplification of SAWs in layered structures (semiconductor–piezoelectric) was observed almost simultaneously by K. Ioshida and M. Yamanishi [30] in a structure consisting of a Ge plate over a piezoceramic plate; by J.H. Collins, K.M. Lakin, C.F. Quate, and J.H. Shaw [31] in a structure consisting of a Si film on a $LiNbO_3$ crystal (see also [32] by Yu.V. Gulyaev, A.M. Kmita, I.M. Kotelyansky, A.V. Medved, and Sh.S. Tursunov [33]); and on a

[2] There were two patents, by W.S. Mortley (1963) and J.H. Rowen (1963), both containing the idea of an interdigital transducer (IDT). In 1963, W.E. Newell applied for a patent at the Westinghouse company for a resonator structure of the type of the IDT proposed by White and Voltmer for the generation and reception of SAW. However, he was refused. But at that time, R. White and F. Voltmer did not know of the existence of these patents, so their work is indeed the first open publication on the IDT.

monolithic-structure CdS film on a Ge substrate by L.A. Coldren, and G.S. Kino [34]. An effective acoustoelectronic SAW amplifier was demonstrated later by Yu.V. Gulyaev, I.M. Kotelyansky, A.V. Medved, and R.A. Mishkinis [35]. The detailed theory of SAW absorption and amplification and of the acoustoelectric effect in piezoelectric semiconductors and in layered piezoelectric–semiconductor structures was developed in [36]. The suggestion of the interdigital transducer (IDT) is indeed a very important key point, and today IDTs in different variations are a main part of all SAW devices.

The works [27] and [28] were the first publications in which it was suggested to use SAWs for information signal processing. The main principle of SAW devices is that the SAW propagates along the free surface of the medium and, thanks to this, is accessible at every point along its path. Thus, it is possible to contact it, influence it, transform it, amplify it, etc., along all the path of its propagation.

Precisely this "accessibility" of SAWs allowed one to consider SAW devices as an effective realization of the so-called model of a "transversal filter." This was a concept of signal processing suggested in 1940 by H.E. Kallmann. Today, SAW devices indeed demonstrate the full scale of the possibilities of this concept for the processing of both analog and digital signals.

The first effective IDT, and the one that was closest to the idea of transversal-filter construction, was suggested (this is another key point) in 1971 independently by A.V. Kovalev and I.B. Yakovkin [37] and by R.F. Tancrell and M.G. Holland [38]. They suggested, for the formation of the required amplitude–frequency curve (AFC), using the variation of the overlapping lengths of the electrodes: the so-called "apodization." In this construction the electric fields under the electrodes, which excite the SAW, can be approximated by δ sources. It is obvious that the amplitude of each δ source (each pair of electrodes of the IDT) is determined by the overlapping length of neighboring electrodes. Calculations show that, in this "ideal" case, the amplitude–frequency curve (AFC) of such an IDT is the Fourier transform of the overlapping of the electrodes as a function of coordinates along the path of SAW propagation.

Unfortunately, this is all true only theoretically, i.e., in the ideal case. All formulas for the ideal transversal filter calculate the characteristics of the SAW IDT only in the first approximation, since they do not take into account the real physical processes that occur when the SAW propagates along the surface of the piezoelectric with metallic electrodes on it. These physical processes lead to "secondary effects" that can drastically change the characteristics of the IDT.

What are these "secondary effects"? Let us list most of them in arbitrary order:

(1) reflections of the SAW inside the transducers themselves;

(2) influence of the mass and electrical conductivity of the electrodes and, hence, of the shape of the electrode structure of the transducer;

(3) influence of diffraction (especially in the region of small overlapping of the electrodes), absorption, and dispersion of the SAW;

(4) influence of the reflection of the SAW from both the input and output transducers;

(5) influence of the impedances of both the source and the load;

(6) influence of parasitic bulk acoustic waves that are generated in transducer structures together with SAWs;

(7) influence of the direct transit of electromagnetic waves via the substrate, the package, or through the external electric circuits.

To avoid the parasitic influence of these "secondary effects," various constructive solutions were suggested, many of which were principal key points. Thus, in 1973, E.G.S. Paige and coworkers [39] suggested a so-called "multistrip coupler." This is a construction that, with the help of metallic strips on the surface, transmits the SAW electric potentials to the other part of the surface and excites the SAW in another channel. Some of the "secondary effects" mentioned above can be reduced by this operation.

To correct the most deleterious "secondary effects" of the AFC shape, namely, (1), (2), (3), and (4), in 1977 Yu.V. Gulyaev, A.M. Kmita, and A.S. Bagdasarian [40] and, independently, S.C. Malocha and B.J. Hunsinger [41] suggested the so-called "capacitive weighting of electrodes." In this construction, the main array of SAW-emitting electrodes (that have uniform overlapping) is connected with complement electrode arrays with apodization by capacitive links. Only these complement (additional) electrode arrays have electrical (ohmic) contacts with the bus lines. Besides reducing the "secondary effects" mentioned above, this construction is much more flexible for the formation of any arbitrary complex amplitude–frequency curve and phase–frequency curve, since there are more places where such characteristics may be formed. One of the most important advantages of the capacitive tap weighted IDT is that both the input and output IDTs may have capacitive apodization, meaning that their AFCs may be multiplied by each other. This can drastically improve the quality of the filter. In the case of usual apodization inside the SAW propagation channel, only one of the IDTs can be apodized, while another IDT should be a wideband one without apodization. This is connected with the distortion of the wave front due to apodization inside the channel, which prevents multiplication of the AFCs.

In 1969, K. Ingebrigtsen [42] suggested a very convenient phenomenological method for the calculation of the characteristics of SAW devices via the "effective dielectric constant" of a piezoelectric semispace by using the measurements of the velocity of SAWs on

free and metallized surfaces. This "Ingebrigtsen method" is widely used even in complicated cases.

E. Ash in 1967 [43] suggested the so-called topographic SAW waveguide, and D.L. White, also in 1967, suggested the "strip" waveguide [44], both of which are used to concentrate acoustic energy along certain paths and which are very important elements of SAW devices. In 1970, E. Ash proposed the SAW resonators and resonator filters [45] that today are widely used in SAW devices.

Most SAW devices use surface acoustic Rayleigh waves [46], which were first considered by Lord Rayleigh in 1885 in connection with problems posed by earthquakes. These waves have no dispersion both in nonpiezoelectric and in piezoelectric materials. They propagate in a surface layer that has a thickness on the order of one wavelength. Thus, for frequencies higher than 1 GHz, these waves propagate in a layer less than 1 micron thick, and surface treatment can unfortunately spoil the properties of a material in this layer. In 1968, J. Bleustein [47] and, independently, Yu.V. Gulyaev [48] suggested a new type of nondispersive SAW, namely, pure shear surface acoustic waves in piezoelectric materials.[3]

They considered piezoelectric crystals of hexagonal symmetry (CdS type) where the wave propagates perpendicular to the C axis with mechanical displacement along the C axis. The satisfaction of stress-free boundary conditions is here due to the exact compensation at the surface by the stresses produced by mechanical displacement and by the piezoeffect. The surface character of this wave, which in textbooks is called the Bleustein–Gulyaev (BG) wave, is due to the reduction of the "stiffened" piezoelectric constant near the surface and the subsequent reduction of the shear bulk acoustic wave velocity in the surface layer (see [50]). In 1970, J. Kerber and R.F. Vogel published a paper (IEE Trans. Sonics Ultrason., Vol. 19, No. 3) entitled "Generalized Bleustein Modes." This theoretical paper presented the general conditions for the existence of shear SH surface waves and specified the crystal cuts where they are possible. This was a generalization of the previous publications [27, 28] and provided a list of the types of piezoelectric crystals, in which such SH surface waves can exist. The first experimental observation of BG waves was done by P. Tournois, C. Maerfeld, and F. Jires [51] in 1970 and, independently, by A.I. Morozov and M.A. Zemlyanitzyn [52] also in 1970.

The broad investigations of physical phenomena connected to the interaction of SAWs with electric fields and electrons in piezoelectric dielectrics and semiconductors and in layered piezoelectric–semiconductor structures, carried out from 1970s to the 1990s in Europe, the United States, the USSR, Japan, and in other countries led to the rapid development of SAW devices and their use in various radioelectronic systems for information processing and communications (see the review paper by D. Morgan [53]). In 1974, five European scientists, namely, E. Ash, J. Collins, Yu. Gulyaev, K. Ingebrigtsen, and E. Paige were awarded the Hewlett-Packard European Physical Society Prize for the development of the physical foundations of SAW devices.

Parallel to the development of the SAW devices in the sixties and seventies, detailed investigations of acoustic-wave (AW) propagation in conducting materials (semiconductors and metals) and of AW interaction with free electrons were carried out. Many new effects were discovered and new methods were developed. Among the most important achievements in this field were new methods of characterizing materials—by their mechanical, electrical, magnetic, thermal, etc., properties—with the help of acoustic waves. These are in the works of F.S. Hickernell and his coworkers [54–57], Slobodnik (see his book [58]), S.N. Ivanov *et al.* [59], A.I. Morozov [60], and many others.

Another achievement was the theoretical prediction and experimental observation of new transport phenomena connected with the dragging of electrons by acoustic waves. One of these phenomena had already been described, namely, the acoustoelectric (AE) effect discovered in 1953 by R. Parmenter. Another effect of this type is the acoustomagnetoelectric (AME) effect: the excitation, by an acoustic wave in a specimen placed in a magnetic field that is transverse to the AW propagation direction, of an electric current (or voltage if the specimen is open-circuited) perpendicular both to the magnetic field and to the AW propagation direction. In the case of a bipolar semiconductor, the AME effect was discovered by A.A. Grinberg and N.I. Kramer [61], and it is due to the dragging of both electrons and holes by the AW and the resulting Hall effect on the acoustoelectric currents related to the electron and holes. In a monopolar semiconductor, AME effects were theoretically predicted by E.M. Epstein and Yu.V. Gulyaev [62] and experimentally observed by A. Korolyuk and N. Roy [63].

In the monopolar semiconductor, the AME effect is connected with the fact that, due to the energy dependence of the electron momentum relaxation time, electrons of different energies are dragged by the AW with different forces. In the case of a short-circuited specimen, the AME current is just the Hall current with a different ("acoustoelectric") Hall coefficient. In the case of an open-circuited specimen, there appears a compensating electric field that drags the electrons in the opposite direction with the same force, i.e., independently of electron energy. As a result, two "partial" electric currents, equal in size and opposite in direction, in which the average energy of the electrons and, hence, their

[3] Bleustein's paper appeared in December 1968 and Gulyaev's paper in January 1969, but the latter was submitted 20 days before Bleustein's. Both works were carried out independently. It should be noted that a shear SAW whose surface character is provided by the piezoeffect was considered earlier by M. Kaganov and S. Sklovskaya [49], but they considered piezocrystals of cubic symmetry, in which such waves cannot exist.

mobilities are different, appear in the specimen in orientations transverse to the magnetic field and AW propagation directions. Thus, the Hall effects due to these "partial" currents do not compensate each other, and a "differential" acoustomagnetoelectric field appears, which depends on the momentum relaxation mechanism of electrons. This effect is a tool for the study of electron scattering mechanisms in solids. The presence of these partial current leads to other "acoustic" electron-transport phenomena: the acoustothermal effect [64], the acoustomagnetothermal effect [65], the "acoustic" Ettingshausen and Peltier effects [66], etc., in which the moving force is the dragging of electrons by the acoustic wave.

Surface acoustic waves produce both longitudinal and transverse (perpendicular to the surface) acoustoelectric effects. The transverse AE effect, which is due to SAW and which was theoretically predicted by Yu.V. Gulyaev *et al.* [36] and experimentally observed by A.M. Kmita and A.V. Medved [67], is a basis for the creation of effective SAW convolvers, correlators, and image-reading devices [68]. The interaction of SAWs with a two-dimensional electron gas in thin films and surface layers was first considered in 1976 by Yu.V. Gulyaev and R. Gasparian [69].

The nonlinear theory of the interaction of an AW of arbitrarily large amplitude with electrons in semiconductors was developed at the end of the sixties. In the works of Y. Abe [70], I.R.A. Beale [71], and P.E. Zilberman [72], stationary AWs that are due to nonlinear effects are considered. P.K. Tien [73] numerically calculated some of the nonlinear effects of AW interaction with electrons for certain cases.

The analytical nonlinear theory of the interaction of AWs of arbitrary large amplitudes with electrons in semiconductors was developed in 1970 by Yu.V. Gulyaev [74].

The analytical nonlinear theory of AW propagation for the nonclassical case (the so-called "momentum" nonlinearity) was developed by P.E. Zilberman [75] in 1971 and experimentally verified that same year by S.N. Ivanov, I.M. Kotelyansky, G.D. Mansfeld, and E.N. Khazanov [76].

In the late 1960s, a nonlinear theory of interaction between acoustic waves of arbitrary high amplitude and electrons in semiconductors was developed. Abe [70], Beale [71], and Zilberman [72] considered stationary acoustic waves caused by nonlinear effects. Tien [73] numerically calculated some of the nonlinear effects of the interaction of acoustic waves with electrons for a number of specific cases.

A detailed experimental study of different mechanisms of acoustoelectronic nonlinearity was carried out using the idea of the multiflight generation and amplification of high-intensity acoustic waves, which was formulated and developed by A.S. Bugaev, Yu.V. Gulyaev, and G.D. Mansfeld in 1978.

We have mentioned here only a few achievements in the early stage of development of acoustoelectronics (before the beginning of the seventies) that, in our opinion, are important contributions to solid-state physics and electronics and that determined the progress of this branch of science and technology for years afterward. More than 40 years have already passed since the first works on SAW applications in electronics appeared. During this period, many new physical works have been carried out, thousands of scientific articles and more than ten large monographs have been published, hundreds of patents have been registered, and many acoustoelectronic devices have been developed.

In this short report, it is not possible to describe everything that has been done, as well as to mention the names of all the scientists and engineers who have contributed to this field. Nevertheless, we dare to give a list (of course, far from complete) of the names of those who, in our opinion, have made the main contributions to the early stage of the development of acoustoelectronics (in the 1960s and the beginning of the 1970s). In alphabetical order, the list is as follows:

USA: R. Adler, B. Auld, J. Bleustein, L. Coldren, C. Hartman, F. Hickernell, M. Holland, A. Hutson, S. Joshi, B. Khuri-Yakub, G. Kino, K. Lakin, D. Malocha, H. Mathews, M. Pomerantz, C. Quate, J. Shaw, A. Slobodnik, H. Smith, R. Smith, H. Spector, R. Tankrell, P. Tien, H. Tiersten, C. Wang, G. Weinreich, D. White, R. White, H. Whitehouse,….

USSR (Russia): V. Anisimkin, M. Balakirev, V. Bonch-Bruevich, S. Bogdanov, A. Bugaev, A. Chaban, L. Chernozatonsky, E. Epstein, Yu. Galperin, A. Ganapolsky, M. Grigor'ev, Yu. Gulyaev, V. Gurevich, Yu. Ilisavsky, S. Ivanov, M. Kaganov, A. Kmita, I. Kotelyansky, I. Kucherov, V. Lemanov, V. Levin, V. Lyamov, R. Maev, G. Mansfeld, A. Medved, A. Morozov, V. Plessky, V. Proklov, V. Pustovoit, V. Shevchik, N. Sinitzyn, Yu. Solodov, K. Tolpygo, I. Victorov, I. Yakovkin, P. Zilberman, Yu. Zyuryukin,….

Japan: N. Chubachi, K. Inoue, S. Inuishi, K. Ioshida, N. Mikoshiba, K. Nakamura, E. Sezawa, K. Shibayama, H. Shimizu, T. Shiosaki, K. Tsubouchi, M. Yamanishi, K. Yamanouchi….

UK: E. Ash, J. Collins, R. De la Rue, J. Maines, D. Morgan, E. Paige, K. Wilkinson,….

France: E. Dieulesaint, J. Henaff, F. Jires, C. Maerfeld, G. Quentin, P. Tournois,….

Norway: K. Blotekjaer, H. Engan, K. Ingebrigtsen,….

Canada: E. Adler, C. Campbell, G. Farnell,….

Germany: W. Buff, K. Dransfeld, M. Weinacht,….

Poland: E. Daniski, S. Kalissky, M. Shustakowsky,….

Many younger scientists and engineers are now very well-known names in SAW devices, physics, and technology. One may find references to works published in the 1970s–1990s in the excellent historical review by D. Morgan [53].

Today, many practical applications of acoustoelectronic phenomena are connected with SAW devices, including

—bandpass filters,

—dispersive filters,

—dispersive delay line,

—devices for coded-signal processing (coders and decoders),

—fast Fourier transformers,

—convolvers and correlators,

—amplifiers and generators,

—digital Nyquist filters,

—frequency synthesizers, etc.

Various SAW bandpass filters, due to their unique characteristics in comparison with other analogous devices, are now used in practically all radioelectronic systems for analog and digital information processing and form more than 90% of the world market in acoustoelectronic devices. These are filters for television, stereo radio broadcasting, audio and videotape recorders, compact disc recorders and players and, more and more in recent years, for radio telephones and cellular telephone networks. Of course, there are many special applications for acoustoelectronic devices in military systems. In accordance with data published in 2003, the production volumes of various kinds of filters are as follows:

—traditional TV filters for intermediate frequencies for different standards—over 2 billion pieces per year;

—filters for audio devices, mainly for automobile FM radios, frequency synthesizers—billions of pieces per year;

—filters for cellular phones: radio frequency filters for frequencies up to 2.5 GHz with bandwidths of 25 MHz (GSM), 34 MHz (EGSM), and 75 MHz (TCM)—over 2 billion pieces per year;

—IF filters for frequencies of 70–450 MHz with the bandwidths 300 kHz (GSM), 1.25 MHz (CDMA), and 5 MHz (WCDMA)—over 2 billion pieces per year.

The total production of acoustoelectronic devices based on SAWs and bulk AWs (resonators and vibrators) in 2003 was approximately more than 10 billion dollars.

The main producers of SAW devices are Murata, Kyoto Ceramics, Samsung, LG, Fujitsu, Hitachi, NEC, Tai SAW, Epcos, SAWTEK, Thomson, Vectron, CTS (Motorola), RF Monolithics, Morion, etc.

The production of acoustoelectronic devices using bulk AW high-precision quartz oscillators, quartz resonators and vibrators (OCXO)- and bulk AW filters is now about 2 billion pieces per year (~3 billion dollars) for:

—watches;

—radio telecommunications and telephony;

—navigation and positioning (GPS);

—control and measurements technique;

—emergency services (Kospas, SARSAT);

—rocket and space techniques.

The main producers are Conning Frequency Control, CQE, C-MAC, Piezo Technology, Frequency Electronics, Morion, etc.

Due to their unique features, small size, mechanical durability, reliability, tuning-free operation, good temperature stability, and ability to work in aggressive media under conditions of strong radiation and high (up to 700°C) temperatures, acoustoelectronic devices will always have their proper niche in electronics of the twenty-first century.

The scientific aspects of acoustoelectronics continue to develop. Though not pretending to predict all but the most important tendencies in the development of future acoustoelectronics in the twenty-first century, we nevertheless will make several notes.

First of all, we should mention SAW matched filters, which are now used for the recognition of coded signals. In the twenty-first century, these filters will have very wide applications as various markers for everything from consumer goods to planes, trains, cars, and even the identification of personnel.

Another direction is connected with the use of bulk acoustic waves (BAWs) at very high frequencies (more than 2 GHz), where SAWs are difficult to apply due to high absorption in the surface layer. The amplification of BAWs at a frequency of 9.4 GHz by the supersonic drift of electrons was obtained by M. Pomerantz [77] in 1964 at 4.2 K and by S.N. Ivanov and G.D. Mansfeld [78] in 1969 at 77 K. Today, the high-frequency devices using BAWs are represented by high-quality resonators and resonator filters [79] that are widely used in radioelectronic systems, and their area of application will surely grow.

Next, we should point out SAW applications in sensors and actuators. Today, SAW sensors are already being used for the identification of gases, vapors, and liquids. Recently, many new constructions of SAW sensors with higher sensitivity and selectivity were suggested, which will open new areas for their application, including narcotics identification [80–82].

The fourth direction of acoustoelectronics development in the twenty-first century will definitely be connected with the use of piezoelectric semiconductors or layered piezoelectric–semiconductor structures, as was suggested in [27]. We can point out at least six acoustoelectronic devices that are based on layered piezoelectric–semiconductor structures:

(1) A SAW amplifier employing the supersonic drift of electrons of the TWT type (see [27], as well as [30–35]). The best result of [35] (central frequency 280 MHz, amplification 50 db, noise factor <7, wide band) is comparable to transistor amplifiers and allows one to hope that this amplifier will find its niche, since it has certain advantages that include the complete electrical isolation of output from input.

(2) Another prospective device may be the so-called acoustic injection transistor (AIT) [83], in which the amplification of the signal is perfomed by the conductivity modulation of the space between collector electrodes as a result of the bunching of electrons in the AW generated by an input signal. Several constructions of AITs have been suggested, many of which theoretically may have quite good characteristics.

(3) Examples of devices connected with charge transmission by an AW of sufficiently large amplitude are acoustic analogs of CCDs (charge-coupled devices) [84–86]. The "physical" principle of these "ACCDs" is that electrons may be trapped into potential pits, created by the large-amplitude AW in a piezoelectric material, and transmitted with an AW velocity. Another construction of an ACCD may use the acoustoelectric effect: the charge bunches in technologically created potential pits can be transferred from one pit to another via the dragging of electrons by the AW pulse.

(4) Convolvers and correlators based on the transverse acoustoelectric effect (see [36, 67, 68, 87]) are much more effective than those devices using SAWs in piezodielectrics, since the "electron nonlinearity" in piezosemiconductors is much larger than the "lattice nonlinearity." One may think that, in the twenty-first century, "acoustoelectronic" convolvers and correlators will be widely used for information processing.

(5) Another perspective is the acoustic image-reading device based on the propagation of short acoustic pulses in a layered-structure piezoelectric-photoconducting semiconductor and the transverse acoustoelectric effect. This is an analog of videcon, only instead of electron-beam scanning, acoustic pulses are used.

(6) Acoustic memory devices may be based on the effect of the trapping of secondary electrons produced by an external electron beam pulse in the surface layer of a piezoelectric in accordance with the potential relief of a traveling SAW [88, 89]. It looks like the SAW is "stopped" for a long time (hours or days, depending on the residual conductivity of the piezoelectric). The reading of the "recorded" information is performed by the application to the surface of another short electron beam pulse, which shorts the piezoelectric fields while the existing stresses excite the same SAW, which can then be registered by an output transducer. It should be noted that, recently, the phenomenon of the "stopping" of light [90] in some gases was experimentally observed. One may see that it is an exact analog of the above-mentioned previously observed "stopping" of a SAW [88, 89].

In conclusion, we want to say that the study of the propagation of acoustic waves in various solid-state materials and their interaction with electric and magnetic fields and with elementary excitations in these solids will, no doubt, lead to the discovery of new interesting effects, which, in turn, will bring new breakthroughs in the creation of high-technology devices in the twenty-first century.

REFERENCES

1. W. G. Cady, *Piezoelectricity* (McGraw-Hill, New York, 1946; Inostrannaya Literatura, Moscow, 1949).
2. *Physical Acoustics. Principles and Methods*, Ed. by W. P. Mason (Academic, New York, 1964; Mir, Moscow, 1966), Vol. 1, Part A.
3. B. Auld, *Acoustic Fields and Waves in Solids* (Wiley, New York, 1973).
4. I. G. Shaposhnikov, Zh. Éksp. Teor. Fiz. **11**, 332 (1941).
5. F. V. Shubnikov, *Piezoelectric Devices* (Moscow, 1940).
6. R. H. Parmenter, Phys. Rev. **89**, 990 (1953).
7. H. E. Bommel, Phys. Rev. **96**, 200 (1954).
8. A. B. Pippard, Philos. Mag. **46**, 1104 (1955).
9. G. Weinreich, Phys. Rev. **107**, 317 (1957).
10. I. M. Lifshitz and M. I. Kaganov, Usp. Fiz. Nauk **69**, 419 (1959) [Sov. Phys. Usp. **2**, 831 (1959)].
11. A. R. Hutson, J. H. McFee, and D. L. White, Phys. Rev. Lett. **7**, 237 (1961).
12. K. B. Tolpygo and Z. I. Uritsky, Zh. Éksp. Teor. Fiz. **30**, 929 (1956) [Sov. Phys. JETP **3**, 725 (1956)].
13. G. Weinreich, Phys. Rev. **104**, 321 (1956).
14. D. L. White, J. Appl. Phys. **33**, 2547 (1962).
15. H. N. Spector, Phys. Rev. **127**, 1054 (1962).
16. M. E. Gertzenstein and V. I. Pustovoit, Radiotekh. Élektron. (Moscow) **7**, 1009 (1962).
17. V. L. Gurevich, Sov. Phys. Solid State **4**, 909 (1962).
18. N. Mikoshiba, J. Phys. Soc. Jpn. **15**, 1189 (1962).
19. R. F. Kazarinov and V. G. Skobov, Zh. Éksp. Teor. Fiz. **42**, 910 (1962) [Sov. Phys. JETP **15**, 628 (1962)].
20. A. R. Hutson and D. L. White, J. Appl. Phys. **33**, 40 (1962).
21. V. L. Gurevich, Sov. Phys. Solid State **4**, 1380 (1962).
22. V. L. Gurevich and V. D. Kagan, Sov. Phys. Solid State **4**, 2441 (1962).
23. H. N. Spector, Solid State Phys. **19**, 291 (1966).
24. V. L. Gurevich, Fiz. Tverd. Tela (Leningrad) **10**, 1557 (1968) [Sov. Phys. Solid State **10**, 1233 (1968)].
25. V. I. Pustovoĭt, Usp. Fiz. Nauk **97**, 257 (1969) [Sov. Phys. Usp. **12**, 105 (1970)].
26. R. W. Smith, Phys. Rev. Lett. **9**, 87 (1962).
27. Yu. V. Gulyaev and V. I. Pustovoĭt, Zh. Éksp. Teor. Fiz. **47**, 2251 (1964) [Sov. Phys. JETP **20**, 1508 (1964)].
28. R. M. White and F. Voltmer, Appl. Phys. Lett. **7**, 314 (1965).
29. I. A. Victorov, Sov. Phys. Acoust. **7**, 295 (1961).
30. K. Yoshida and M. Yamanishi, Jpn. J. Appl. Phys. **7**, 1143 (1968).
31. J. H. Collins, K. M. Lakin, C. F. Quate, and J. H. Shaw, Appl. Phys. Lett. **13**, 314 (1968).
32. J. H. Collins, H. M. Gerard, and H. J. Shaw, Appl. Phys. Lett. **13**, 312 (1968); K. M. Lakin and H. J. Shaw, IEEE Trans. Microwave Theory Tech. **17**, 912 (1969).
33. Yu. V. Gulyaev, A. M. Kmita, I. M. Kotelyansky, *et al.*, Fiz. Tverd. Tela (Leningrad) **13**, 1557 (1971) [Sov. Phys. Solid State **13**, 1305 (1971)].
34. L. A. Coldren and G. S. Kino, Appl. Phys. Lett. **23**, 117 (1973).

35. Yu. V. Gulyaev, I. M. Kotelyansky, A. V. Medved, and R. A. Mishkinis, Electron. Lett. **16**, 114 (1980).
36. Yu. V. Gulyaev, A. Yu. Karabanov, A. M. Kmita, *et al.*, Sov. Phys. Solid State **12**, 2595 (1970).
37. A. V. Kovalev and I. B. Yakovkin, Radiotekh. Élektron. (Moscow) **16**, 1521 (1971).
38. R. H. Tancrell and M. G. Holland, Proc. IEEE **59**, 393 (1971).
39. F. G. Marshall and E. G. S. Paige, Electron. Lett. **7**, 460 (1971); F. G. Marshall, C. O. Newton, and E. G. S. Paige, IEEE Trans. Sonics Ultrason. **20**, 124 (1973).
40. Yu. V. Gulyaev, A. M. Kmita, and A. S. Bagdasarian, Sov. Tech. Phys. Lett. **5**, 287 (1979); SU Patent No. 726648; US Patent No. 4,162,415; US Patent No. 4,185,218; GB Patent No. 2,003,353; JP Appl. No. 1,069,686; FR Patent No. 7,821,723; FR Patent No. 8,020,674, DE Patent No. 2,831,584, DE Patent No. 2,831,585.
41. D. C. Malocha and B. J. Hunsinger, IEEE Trans. Sonics Ultrason. **24**, 293 (1977).
42. K. Ingebrigtsen, J. Appl. Phys. **40**, 2681 (1969).
43. E. A. Ash, in *Proceedings of the IEEE Symposium on Microwave Theory and Technology* (Boston, 1967).
44. D. L. White, in *Proceedings of the IEEE Ultrasonic Symposium* (Vancouver, 1967).
45. E. A. Ash, in *Proceedings of the IEEE Symposium on Microwave Theory Technology* (Newport Beach, 1970).
46. Lord Rayleigh, Proc. London Math. Soc. **17**, 4 (1885).
47. J. Bleustein, Appl. Phys. Lett. **13**, 412 (1968).
48. Yu. V. Gulyaev, JETP Lett. **9**, 63 (1969).
49. M. I. Kaganov and S. Sklovskaya, Sov. Phys. Solid State **9** (1967).
50. Yu. V. Gulyaev, *Rep. at the UFFC Symposium* (Seattle, 1995); IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**, 935 (1998).
51. P. Tournois, C. Maerfeld, and A. Jires, Appl. Phys. Lett. (1970).
52. A. I. Morozov and M. I. Zemlyanitsyn, JETP Lett. (1970).
53. D. Morgan, in *Proceedings of IEEE International Frequency Control Symposium* (Pasadena, 1998).
54. F. S. Hickermell and N. Sakiotis, Proc. IEEE **52**, 194 (1969).
55. J. Wasilik and F. S. Hickernell, Appl. Phys. Lett. **24**, 153 (1974).
56. F. S. Hickernell, Int. J. High Speed Electron. Syst. **10**, 603 (2000).
57. F. S. Hickernell and T. S. Hickernell, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42** (3), 410 (1995).
58. A. Slobodnik, R. Delmonico, and E. Conway, *Microwave Acoustics Handbook* (Air Force Cambridge Lab., Bedford, Mass., 1970).
59. S. N. Ivanov, S. F. Akhmetov, I. M. Kotelyansky, and V. V. Medved, Sov. Phys. Solid State **19**, 308 (1977); S. N. Ivanov, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 653 (1992); Yu. V. Gulyaev, S. N. Ivanov, and A. G. Kozorezov, Radiotekh. Élektron. (Moscow) **23**, 2396 (1978).
60. A. I. Morozov and M. A. Zemlyanitsyn, Sov. Phys. Solid State **6**, 2298 (1972); A. I. Morozov, M. A. Zemlyanitsyn, and V. I. Anisimkin, Phys. Status Solidi A **14**, 339 (1974); Phys. Status Solidi A **24**, 381 (1974).
61. A. A. Grinberg and N. I. Kramer, Dokl. Akad. Nauk SSSR **157**, 79 (1964) [Sov. Phys. Dokl. **9**, 552 (1965)].
62. E. M. Epstein and Yu. V. Gulyaev, Sov. Phys. Solid State **9**, 376 (1967).
63. A. Koroluk and N. Roy, Sov. Phys. Solid State **16** (1974).
64. Yu. V. Gulyaev and E. M. Epstein, JETP Lett. **3**, 410 (1966).
65. Yu. V. Gulyaev and E. M. Epstein, Sov. Phys. Solid State **9**, 864 (1967).
66. Yu. V. Gulyaev, Sov. Phys. Solid State **8**, 3366 (1966).
67. A. M. Kmita and A. V. Medved, JETP Lett. **14**, 455 (1971); J. Appl. Phys. **44**, 3034 (1973).
68. W. C. Wang and P. Das, in *Proceedings of IEEE Ultrasonic Symposium* (1972), p. 316.
69. Yu. V. Gulyaev and R. A. Gasparian, Microelectronics **8**, 326 (1979); Surf. Sci. **98**, 553 (1980).
70. Y. Abe, Prog. Theor. Phys. **31**, 956 (1964).
71. I. R. A. Beale, Phys. Rev. **136**, 1761 (1964).
72. P. E. Zilberman, Sov. Phys. Solid State **9**, 309 (1967).
73. P. K. Tien, Phys. Rev. **171**, 970 (1968).
74. Yu. V. Gulyaev, Sov. Phys. Solid State **12**, 415 (1970); IEEE Trans. **415**, 19 (1970).
75. P. E. Zilberman, Sov. Phys. Solid State **5**, 1240 (1971).
76. S. N. Ivanov, I. M. Kotelyansky, G. D. Mansreld, and E. N. Khazsnov, JETP Lett. **13**, 283 (1971).
77. M. Pomerantz, Phys. Rev. Lett. **13**, 308 (1964).
78. S. N. Ivanov and G. D. Mansfeld, Radiotekh. Élektron. (Moscow) **14**, 368 (1969).
79. Yu. V. Gulyaev and G. D. Mansfeld, Radiotekh. Élektron. (Moscow) **8**, 1529 (2003).
80. V. I. Anisimkin, R. G. Kryshtal, A. V. Medved, *et al.*, Electron. Lett. **34**, 1360 (2000).
81. I. V. Anisimkin, V. I. Anisimkin, and Yu. V. Gulyaev, in *Proceedings of IEEE Ultrasonic Symposium* (Puerto Rico, 2000), p. 713.
82. I. V. Anisimkin, in *Proceedings of IEEE Ultrasonic Symposium* (Hawaii, 2003), p. 1326.
83. Yu. V. Gulyaev, SU Patent (1971); Yu. V. Gulyaev, G. D. Mansfeld, and G. A. Orlova, Electron. Lett. **17** (12) (1981).
84. Yu. V. Gulyaev, SU Patent (November 1971).
85. A. Siegert, Austrian Patent (December 1971).
86. R. L. Miller, C. E. Northwic, and D. S. Bailey, *Acoustic Charge Transport: Device Technology and Application* (Artech, Boston, 1992).
87. G. D. Mansfeld, in *Proceedings of IEEE Ultrasonic Symposium* (Hawaii, 2003).
88. P. S. Voronov, Z. S. Chernov, *et al.*, Pis'ma Zh. Tekh. Fiz. (1971).
89. A. G. Bert, B. Epstei, and G. Kantorovicz, IEEE Trans. Microwave Theory Tech. **21**, 255 (1973).
90. M. D. Lukin, M. Fleischhauer, R. Cote, *et al.*, Phys. Rev. Lett. **87**, 037901 (2001).

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 89–94. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 111–116.

A Monopole–Dipole Resonance Absorber in a Narrow Waveguide

N. G. Kanev and M. A. Mironov

Andreev Acoustics Institute, Russian Academy of Sciences, ul. Shvernika 4, Moscow, 117036 Russia

e-mail: mironov@akin.ru

Received June 23, 2004

Abstract—A sound absorber in a narrow waveguide is considered. The absorber consists of one monopole and one dipole resonator placed in a narrow pipe. The optimum parameters of the resonators that provide for the maximum absorption of acoustic power are determined. Results of an experimental study of a two-resonator absorbing system are presented. A 95% absorption is achieved. © *2005 Pleiades Publishing, Inc.*

Using a nondissipative resonator of a monopole (e.g., a Helmholtz resonator [1]) or dipole [2] type, it is possible to provide for a total reflection of sound propagating in a narrow waveguide. However, when used as an absorber of sound, a resonator shows a much lower efficiency. A single resonator with an optimal loss is capable of absorbing only half the incident acoustic power. A complete absorption of sound is possible with a combination of resonators. Lapin [3] proposed the use of two Helmholtz resonators, one of which provides a total reflection of the incident wave and the other, placed at the pressure antinode, completely absorbs the power. The distance between the resonators should be equal to an odd number of quarter-wavelengths. A complete absorption is also possible when two resonators, namely, a monopole resonator and a dipole one, tuned to the same frequency, are positioned in one waveguide cross section [4]. The resonators operate independently, so that each of them absorbs half the incident acoustic power. Note that a system of monopoles and dipoles that absorbs a sound wave was considered years ago [5]. It was shown that a planar array of monopoles and dipoles arranged in a staggered order so that the distances between them were small compared to the wavelength completely absorbs a plane wave incident on the array under a certain angle.

In this paper, we present the results of an experiment with two resonators, namely, a monopole resonator and a dipole one, placed in a narrow pipe. For some reason, the resonators were positioned not in one cross section of the pipe (as in [4]) but at a small (compared to the wavelength) distance along the pipe axis. In this regard, we present a theoretical generalization of the problem of sound absorption by two resonators, which was considered in [4]. We theoretically investigate the dependence of the absorption efficiency on the distance between the resonators and on their resonance frequency mismatch and calculate the bandwidth within which the absorption is sufficiently effective.

If two closely spaced resonators are used to provide for the sound absorption in a pipe, it is necessary that they be of different types. If one of the resonators, e.g., the monopole one, is used as a reflector, a particle velocity antinode is formed near it on the side of the incident wave. To absorb the incident sound, it is necessary to place a dipole resonator with an optimal absorption at this antinode. Analogously, if the dipole resonator is used as a reflector, a pressure antinode is formed near it and, hence, an absorbing monopole should be placed there. Below, we study an absorbing system that consists of one monopole and one dipole resonator with the monopole resonator serving as a reflector.

A harmonic wave of unit amplitude with a frequency ω and a wave number k propagates in the positive direction of the x axis of a narrow waveguide (Fig. 1). A dipole resonator is placed at the point $x = 0$, and a monopole resonator is placed at $x = L$. The system of two resonators reflects a wave with an amplitude V and transmits a wave with an amplitude W. Thus, the pressure field to the left of the resonators can be represented as $P_l(x) = \exp(ikx) + V\exp(-ikx)$ and the pressure field to the right, as $P_r(x) = W\exp(ikx)$. The field between the two resonators is also represented as a sum of two waves traveling in opposite directions: $P_1(x) = A\exp(ikx) + B\exp(-ikx)$. The particle velocity fields are represented in a similar way: $U_l(x) = (\exp(ikx) - V\exp(-ikx))/\rho c$, $U_r(x) = W\exp(ikx)/\rho c$, and $U_1(x) =$

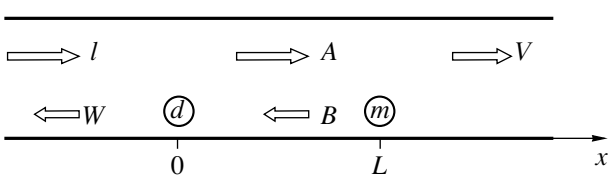


Fig. 1. Monopole (m) and dipole (d) resonators in a waveguide.

1063-7710/05/5101-0089$26.00 © 2005 Pleiades Publishing, Inc.

$(A\exp(ikx) - B\exp(-ikx))/\rho c$, where ρ is the density of the medium and c is the sound velocity in it.

For the cross section in which the dipole is positioned, two boundary conditions are satisfied: (i) the velocities to the right and to the left of the dipole are equal and (ii) the pressure drop is equal to the dipole strength f divided by the area of the waveguide cross section S. These conditions are expressed as

$$U_l(0) = U_1(0), \tag{1}$$

$$P_l(0) - P_1(0) = f/S. \tag{2}$$

For the cross section in which the monopole is positioned, two additional boundary conditions should be satisfied: (i) the pressures to the left and to the right of the monopole are equal and (ii) the particle velocity drop is equal to the monopole volume velocity q divided by S. These conditions have the form

$$U_1(L) - U_r(L) = q/S, \tag{3}$$

$$P_1(L) = P_r(L). \tag{4}$$

The dipole strength and the monopole volume velocity are determined from the corresponding equations of motion. The equation of motion of the dipole resonator has the form [2]

$$m_1\dot{v}_1 = -\mu(\dot{v}_1 - \dot{u}) - \kappa_1\xi_1 - \gamma_1 v_1, \tag{5}$$

where v_1 and ξ_1 are the velocity and displacement of the dipole, respectively; m_1 is its mass; μ is the associated mass; κ_1 is the coefficient of elasticity; γ_1 is the friction factor; and u is the velocity of the medium near the dipole. The first term on the right-hand side of Eq. (5) is the force acting on the dipole from the side of the medium. This quantity with the opposite sign is the dipole strength. From boundary condition (1), we obtain a expression for the velocity of the medium near the dipole: $u = (1 - V)/\rho c$. Thus, the dipole strength can be expressed through the sound field as

$$f = \frac{i\omega\mu}{\rho c}\frac{i\omega\mu + D_1}{D_1}(1 - V), \tag{6}$$

where $D_1 = -i\omega(m_1 + \mu) + \gamma_1 + \dfrac{\kappa_1}{-i\omega}$.

The equation of motion of the monopole resonator has the form [1]

$$m_2\dot{v}_2 = -p\sigma - \kappa_2\xi_2 - \gamma_2 v_2, \tag{7}$$

where v_2 and ξ_2 are the velocity and displacement of air in the resonator neck, m_2 is the mass of air in the neck, σ is the area of the neck cross section, κ_2 is the elasticity of air inside the resonator, γ_2 is the friction factor, and p is the pressure near the resonator. From boundary condition (4), it follows that the pressure near the monopole is equal to $p = W\exp(ikL)$. Thus, the monopole volume velocity is expressed through the sound field as

$$q = -\frac{\sigma^2}{D_2}W\exp(ikL), \tag{8}$$

where $D_2 = -i\omega m_2 + \gamma_2 + \dfrac{\kappa_2}{-i\omega}$.

With allowance for Eqs. (6) and (8), set of boundary conditions (1)–(4) can be represented in the form

$$\begin{gathered} A - B + V = 1, \\ A + B + (M_1 - 1)V = M_1 + 1, \\ \exp(ikL)A + \exp(-ikL)B - \exp(ikL)W = 0, \\ \exp(ikL)A - \exp(-ikL)B - (M_2 + 1)\exp(ikL)W = 0. \end{gathered} \tag{9}$$

Here, we have introduced the notations

$$M_1 = \frac{i\omega\mu}{\rho cS}\frac{i\omega\mu + D_1}{D_1}, \quad M_2 = -\frac{\sigma^2\rho c}{SD_2}. \tag{10}$$

The physical meaning of the quantities M_1 and M_2 can be understood from the comparison of Eqs. (6) and (8) with Eqs. (10). The dipole impedance is equal to ρcM_1S, and, at the resonance frequency, it tends to infinity with decreasing loss. Hence, at the resonance frequency, the dipole is equivalent to a hard wall. The monopole impedance is equal to $\sigma\rho c/M_2$, and, at the resonance frequency, it tends to zero. Hence, at the resonance frequency, the monopole is equivalent to a soft boundary.

Solving set of equations (9), we obtain the expressions for the coefficients of reflection V and transmission W:

$$V = -\frac{\dfrac{M_1}{2} + \dfrac{M_2}{2}\exp(2ikL) + \dfrac{M_1M_2}{4}[1 + \exp(2ikL)]}{1 - \dfrac{M_1}{2} + \dfrac{M_2}{2} - \dfrac{M_1M_2}{4}[1 + \exp(2ikL)]}, \tag{11}$$

$$W = \frac{1}{1 - \dfrac{M_1}{2} + \dfrac{M_2}{2} - \dfrac{M_1M_2}{4}(1 + \exp(2ikL))}. \tag{12}$$

A complete absorption of the sound wave incident on the system means the absence of transmitted and reflected waves. From Eq. (12) it follows that, for the absence of any transmitted field, one of the following conditions should be satisfied: $M_1 = \infty$ ($D_1 = 0$) or $M_2 = \infty$ ($D_2 = 0$). This means that at least one of the resonators should be nondissipative, i.e., should reflect the whole wave at the resonance frequency. Since the monopole is positioned behind the dipole along the direction of the wave incidence (Fig. 1), it is the monopole that should reflect the sound wave at the corre-

sponding resonance frequency $\omega_2 = \sqrt{\kappa_2/m_2}$. Hence, the friction factor of the monopole γ_2 should be equal to zero. Under the condition $M_2 = \infty$, expression (11) for the reflection coefficient takes the form

$$V = -\frac{\exp(2ikL) + \frac{M_1}{2}[1 + \exp(2ikL)]}{1 - \frac{M_1}{2}[1 + \exp(2ikL)]}. \quad (13)$$

For the reflected field to be equal to zero, it is necessary that the numerator in Eq. (13) be equal to zero; i.e., the following condition should be satisfied:

$$M_1 = -\frac{2\exp(2ikL)}{1 + \exp(2ikL)} = -1 - i\frac{\sin(2kL)}{1 + \cos(2kL)}. \quad (14)$$

Equating the real and imaginary parts of Eq. (14) and assuming that $kL \ll 1$ (the size of the absorber is small), we arrive at a system of two equations that determines the necessary parameters of the dipole resonator:

$$\frac{\omega_2^2\mu^2\gamma_1}{\rho cS\left[\omega_2^2(m_1 + \mu)^2\left(1 - \frac{\omega_1^2}{\omega_2^2}\right)^2 + \gamma_1^2\right]} = 1,$$
$$\frac{\omega_2\mu}{\rho cS}\left[\frac{\omega_2^2\mu^2(m_1 + \mu)\left(1 - \frac{\omega_1^2}{\omega_2^2}\right)}{\omega_2^2(m_1 + \mu)^2\left(1 - \frac{\omega_1^2}{\omega_2^2}\right) + \gamma_1^2} + 1\right] = -k_2L, \quad (15)$$

where $k_2 = \frac{\omega_2}{c}$ and $\omega_1 = \sqrt{\frac{\kappa_1}{m_1 + \mu}}$ is the resonance frequency of the dipole. Solving Eqs. (15) for the resonance frequency ω_1 and the friction factor γ_1 of the dipole, we obtain

$$\omega_1 = \omega_2\sqrt{1 - \frac{\omega_2^2\mu^2}{(m_1 + \mu)(\mu + \rho LS)\left[\omega_2^2 + \left(\frac{\rho cS}{\mu + \rho LS}\right)^2\right]}}, \quad (16)$$
$$\gamma_1 = \frac{\omega_2^2\mu^2}{\rho cS\left[1 + \left(\frac{\omega_2(\mu + \rho SL)}{\rho cS}\right)^2\right]}.$$

The quantity ρSL represents the air mass in the waveguide between the two resonators, and, for a small-size dipole, we can assume that $\mu \ll \rho SL$. Then, Eqs. (16) take the form

$$\omega_1 = \omega_2\sqrt{1 - \frac{\omega_2\mu}{\rho cS}\frac{\mu}{m_1 + \mu}k_2L},$$
$$\gamma_1 = \frac{\omega_2^2m^2}{\rho cS(1 + (k_2L)^2)}. \quad (17)$$

Remember that the frequency corresponding to the zero value of the transmission coefficient is equal to the resonance frequency of the monopole ω_2. According to Eqs. (16) and (17), for a complete absorption of the incident wave, the resonance frequency of the dipole should be lower than the resonance frequency of the monopole. The difference in the resonance frequencies is proportional to the distance L. The friction factor at which the complete absorption takes place exhibits a weaker dependence on L.

When the resonance frequencies of the dipole and the monopole coincide, a complete absorption is possible if the dipole is placed precisely at the antinode of the particle velocity of the standing wave and if the condition $M_1 = -1$ is satisfied. In the case of a compact absorber (the wave distance between the resonators is small but not equal to zero), the latter condition is not satisfied and part of the sound wave is reflected. In the case of coincident resonance frequencies of the dipole and monopole, i.e., $\omega_1 = \omega_2 = \omega_0$, from Eqs. (10) and (13) it follows that, for the maximum absorption, the friction factor of the dipole should be

$$\gamma_1 = \frac{\omega_0^2\mu^2}{\rho cS}. \quad (18)$$

The corresponding minimum coefficient of reflection from the system of resonators is

$$V = -ikL/2. \quad (19)$$

In the limit $L \longrightarrow 0$, we obtain a complete absorption, which is the same result as that obtained in [4]. The difference here from [4] consists in that, in our case, the absorption is provided by only the dipole, while in the geometry considered in [4], with the monopole and dipole lying in the same cross section, the absorption is equally distributed between the monopole and the dipole. Actually, the distance between the monopole and the dipole can be reduced to a value at which the interaction of the resonators through the near field (the mutual associated mass) can still be ignored.

In addition to the maximum value of the absorption coefficient, an important parameter characterizing the efficiency of the absorbing system is the frequency bandwidth within which the absorption coefficient

$$\alpha = 1 - |V|^2 - |W|^2 \quad (20)$$

remains smaller than a certain value α_0. Determining the bandwidth, we assume that the dipole and monopole resonators have the same resonance frequency ω_0

and that the friction coefficient of the dipole is given by Eq. (18). This assumption does not practically affect the result, because the optimal eigenfrequency of the dipole differs little from the eigenfrequency of the monopole (see Eq. (13)). This difference is significant only for the maximum value of the absorption coefficient.

When $kL \ll 1$, Eqs. (11) and (12) take the form

$$V = -\frac{M_1 + M_2 + M_1M_2}{2 - M_1 + M_2 - M_1M_2},$$

$$W = \frac{2}{2 - M_1 + M_2 - M_1M_2}.$$

For small deviations of $M_{1,2}$ from the optimum values $M_1 = -1$ and $M_2 = \infty$, the formulas given above can be approximately represented as

$$V = -\frac{1 + M_1}{1 - M_1}, \quad W = \frac{1}{M_2}. \tag{21}$$

It is convenient to introduce the dimensionless parameter x that determines the deviation of the frequency from the resonance one: $\omega = \omega_0(1 + x)$. Under the assumption that the associated mass of the dipole is much greater than its own mass, i.e., $\mu \gg m_1$, we represent expressions (10) for the case of $x \ll 1$ in the form

$$M_1 = -\frac{\omega_0\mu}{\rho cS}\frac{Q_1}{1 - 2ixQ_1}, \quad M_2 = -\frac{\sigma^2\rho c}{\omega_0 m_2 S}\frac{Q_2}{1 - 2ixQ_2}. \tag{22}$$

Here, $Q_1 = (m_1 + \mu)\omega_0/\gamma_1$ and $Q_2 = m_2\omega_0/\gamma_2$ are the Q factors of the dipole and monopole resonators, respectively.

Substituting Eqs. (22) into Eqs. (21) and then substituting Eqs. (21) into Eq. (20), we obtain the relation between the absorption coefficient and the frequency mismatch x:

$$\alpha \approx 1 - \left(\left(\frac{\rho cS}{\omega_0\mu}\right)^2 + \left(2\frac{S\omega_0 m_2}{\sigma^2\rho c}\right)^2\right)x^2. \tag{23}$$

The factors $\frac{\rho cS}{\omega_0\mu}$ and $\frac{S\omega_0 m_2}{\sigma^2\rho c}$ determine the operating bandwidths of the dipole and monopole resonators, respectively. According to [2], dipole and monopole resonators of the same volume exhibit the same efficiency in terms of the sound reflection in a narrow waveguide. The dipole volume is the volume of the associated mass: $\Omega_1 = \mu/\rho$. The cited paper [2] describes the structure of a dipole resonator that provides for a multiple increase in the associated mass. The monopole volume is the volume of the cavity Ω_2, and its resonance frequency can be expressed through this volume [1]: $\omega_2^2 = \frac{\rho c^2\sigma^2}{m_2\Omega_2}$. Expressing absorption coefficient (23) in terms of the resonator volumes Ω_1 and Ω_2, we determine the bandwidth within which the fraction of the absorbed power is no less than α_0:

$$\Delta\omega = 2x\omega_0 = \frac{\omega_0^2}{cS}\sqrt{\frac{1 - \alpha_0}{\left(\frac{1}{2\Omega_1}\right)^2 + \left(\frac{1}{\Omega_2}\right)^2}}. \tag{24}$$

Factor 2 multiplying Ω_1 in Eq. (24) appears because the dipole plays the role of an absorber rather than a scatterer. From Eq. (24) it follows that the absorption bandwidth increases as the volumes of the dipole and monopole increase. If the total volume of the dipole and monopole is fixed, i.e., $\Omega = \Omega_1 + \Omega_2 = \text{const}$, the maximum absorption bandwidth is achieved when $\Omega_2 = 2\Omega_1 = \frac{2}{3}\Omega$, i.e., when the monopole volume is twice as great as the dipole volume. Let us compare the operating bandwidth $\Delta\omega$ with the difference between the resonance frequencies of the dipole and the monopole $\delta\omega = \omega_1 - \omega_2$ at the optimum tuning given by Eq. (17):

$$\frac{\delta\omega}{\Delta\omega} \approx \frac{\mu}{2\rho}\sqrt{(2\Omega_1)^{-2} + \Omega_2^{-2}}\frac{kL}{\sqrt{1 - \alpha_0}}.$$

When the optimum relation between the dipole and monopole volumes takes place, this ratio is equal to

$$\frac{\delta\omega}{\Delta\omega} \approx \frac{\sqrt{2}}{4}\frac{kL}{\sqrt{1 - \alpha_0}}.$$

From this formula it follows that, when the condition $kL < \frac{4}{\sqrt{2}}\sqrt{1 - \alpha_0}$ is satisfied, the operating bandwidth does not depend on the accuracy of tuning of the two resonators.

Figure 2 illustrates the operation of an absorbing system of two resonators (Fig. 1) tuned to the same frequency ω_0 for different relations between the resonator parameters. The dimensionless distance between the resonators is $k_0L = 0.1$, where $k_0 = \omega_0/c$. The thin lines indicate the amplitude coefficients of transmission (the dashed lines) and reflection (the thin solid lines), and the thick line represents the sum of transmitted and reflected acoustic powers, i.e., the quantity $|W|^2 + |V|^2$. Figure 2a refers to a system in which both resonators have the same operating bandwidth ($\Omega_1 = \Omega_2$). The transmission coefficient is equal to zero, and the reflection coefficient is nonzero according to Eq. (19). However, at the optimum parameters of the resonators, the reflection coefficient V is small and the absorption coefficient α is close to zero but not equal to it. Figure 2b represents the condition under which the operating bandwidth of the dipole is two times smaller ($2\Omega_1 = \Omega_2$). The absorption curve is somewhat narrower, and a

minimum of the reflection coefficient appears at the resonance frequency. The width of the absorption curve noticeably decreases if the operating bandwidth of the monopole is reduced to half its initial value ($\Omega_1 = 2\Omega_2$, Fig. 2c; compare with Fig. 2a).

The absorbing system considered above has been studied experimentally. In the experimental setup, the transmitted and reflected fields are measured by an acoustic interferometer, which has the form of a circular pipe 10 cm in diameter and 2 m in length (Fig. 3) [6]. A sound source is mounted on one end of the interferometer, and a broadband signal is supplied to the source. The other end is filled with a sound-absorbing material to reduce the natural resonances of the pipe. The sample is mounted in the middle of the interferometer, i.e., at a distance of 1 m from the source. The amplitudes of waves *A*, *B*, *C*, and *D* travelling on both sides of the sample are measured by pairs of microphones (see, e.g., [7]). Wave *B* is the sum of wave *A* reflected from the sample and wave *D* transmitted through the sample. The following equality takes place:

$$B = VA + WD. \quad (25)$$

Analogously, wave *D* is a sum of wave *A* transmitted through the sample and wave *D* reflected from the sample. Hence, we have

$$C = WA + VD. \quad (26)$$

Using Eqs. (25) and (26), the reflection and transmission coefficients can be expressed in terms of the measured amplitudes of the four travelling waves that occur in the interferometer:

$$V = \frac{DC - AB}{D^2 - A^2}, \quad W = \frac{DB - AC}{D^2 - A^2}.$$

If the wave transmitted through the sample can be completely absorbed, i.e., if $D = 0$, common relations are valid for the reflection and transmission coefficients: $V = B/A$ and $W = C/A$.

The signals from four microphones are supplied to a four-channel analog-to-digital converter, and the data from the latter are sent to a PC, which performs the signal processing.

The monopole is a cylindrical cavity with a circular hole closed with a membrane (to reduce the resonance frequency) on its one end. The volume of the cavity is 2.2×10^{-4} m^3, and the area of the hole is 7.9×10^{-5} m^2. The dipole has the form of a 2-cm-long pipe with a diameter of 3.7 cm (the cavity volume is 2.15×10^{-5} m^3), and one end of this pipe is closed with a membrane. Here, the membrane plays the role of elasticity. Such a structure is considered in detail in [2]. The parameters of the dipole, namely, the resonance frequency coinciding with the monopole resonance frequency and the optimal loss necessary for the maximum absorption, are provided by adjusting the membrane tension and the length of the pipe. The resonance frequencies of the monopole and the dipole are approximately identical

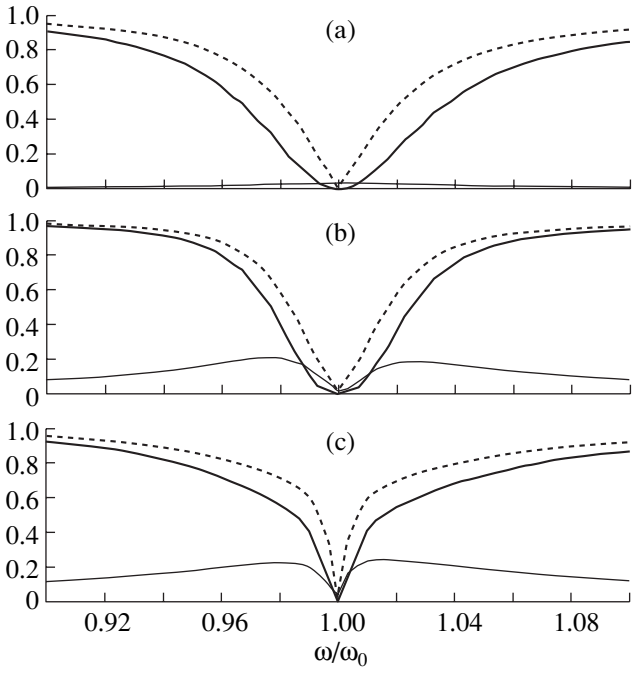


Fig. 2. Coefficients of reflection (thin solid lines), transmission (thin dashed lines), and absorption (thick lines) of the two-resonator system for the following different values of the parameters: (a) $\Omega_1 = \Omega_2 = \Omega$; (b) $\Omega_1 = \Omega/2$, $\Omega_2 = \Omega$; and (c) $\Omega_1 = \Omega$, $\Omega_2 = \Omega/2$.

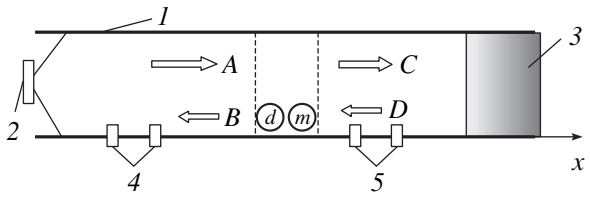


Fig. 3. Schematic representation of the experimental setup: (*1*) waveguide, (*2*) sound source, (*3*) sound absorber, (*4*) pair of microphones measuring the reflected field, and (*5*) pair of microphones measuring the transmitted field.

and equal to 315 Hz. The problem of achieving the necessary resonator mismatch given by Eqs. (17) was discarded, because the theoretical difference $\delta\omega$ for the system under investigation was smaller than 1 Hz. The Q factors of the resonators are determined from their resonance curves and are equal to 90 for the monopole and 30 for the dipole. This value is close to the optimum Q factor at which the friction coefficient takes on the necessary value given by Eqs. (17). The distance between the resonators is 8 cm ($kL = 0.47$).

Figure 4 shows the results of measurements. Two resonators with parameters closest to the optimum ones absorb 95% of the energy of the incident sound wave at the resonance frequency. The frequency dependence of the coefficient $E = 1 - \alpha$ is shown in Fig. 4a. The minimum value of E is –13 dB, and the relative bandwidth within which E does not exceed –10 dB is equal to

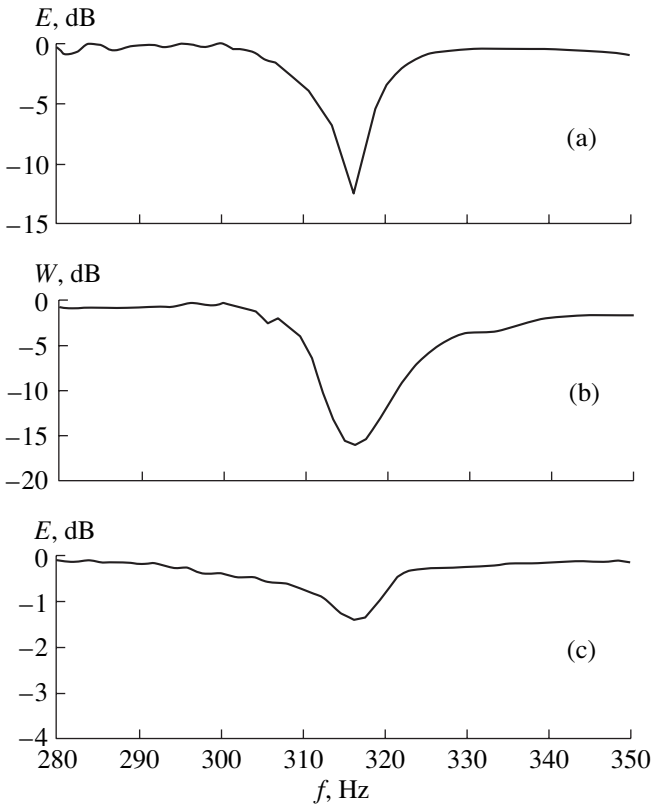


Fig. 4. Results of the experimental study of the resonator operation: (a) absorption coefficient of the monopole–dipole resonator, (b) transmission coefficient of the monopole resonator, and (c) absorption coefficient of the dipole resonator.

0.008. A complete absorption of the incident wave was not achieved, presumably because the dipole loss was not the optimum one.

Figures 4b and 4c show the characteristics of single resonators. The coefficient of transmission (Fig. 4b) through the waveguide cross section, in which the monopole resonator is positioned, is equal to –16 dB at the resonance frequency. The coefficient E (Fig. 4c) for the dipole is characterized by small values and is equal to 1.5 dB at the resonance frequency. The operating bandwidth of the dipole [3] is much smaller than that of the monopole: their ratio is about 1 : 10, i.e., the same as the ratio of their volumes. Therefore, the operating bandwidth of the two-resonator absorbing system is four times smaller than the operating bandwidth of the monopole.

Thus, this study experimentally verifies the possibility of an effective absorption of sound in a narrow pipe by a system of two closely spaced resonators, one of which is a monopole and the other, a dipole. A theory that allows for the optimization of the parameters of the system is developed. Later, we intend to perform a more detailed experimental verification of the theory proposed above, specifically to study the influence of the distance between the resonators and the accuracy of their tuning and dimensions on the absorption bandwidth.

REFERENCES

1. M. A. Isakovich, *General Acoustics* (Nauka, Moscow, 1973) [in Russian].
2. N. G. Kanev and M. A. Mironov, Akust. Zh. **49**, 372 (2003) [Acoust. Phys. **49**, 312 (2003)].
3. A. D. Lapin, Akust. Zh. **49**, 427 (2003) [Acoust. Phys. **49**, 363 (2003)].
4. A. D. Lapin, Akust. Zh. **49**, 855 (2003) [Acoust. Phys. **49**, 731 (2003)].
5. G. D. Malyuzhinets, Tr. Akust. Inst. Akad. Nauk SSSR, No. 15, 7 (1971).
6. M. A. Mironov, I. I. Sizov, A. Ya. Gorenberg, V. S. Solntseva, V. V. Tyutekin, and F. F. Kamenets, *Acoustic Waveguides* (Mosk. Fiz.–Tekh. Inst., Moscow, 2003) [in Russian].
7. A. F. Seybert and D. F. Ross, J. Acoust. Soc. Am. **61**, 1362 (1977).

Translated by E. Golyamina

Acoustical Physics, Vol. 51, No. 1, 2005, pp. 95–104. Translated from Akusticheskiĭ Zhurnal, Vol. 51, No. 1, 2005, pp. 117–127.

Experimental Study of the Scattering of Ultrasound from Wake Flows in a Linearly Stratified Fluid

V. E. Prokhorov and Yu. D. Chashechkin

Institute of Problems of Mechanics, Russian Academy of Sciences, pr. Vernadskogo 101, korp. 1, Moscow, 119526 Russia
e-mail: prohorov@ipmnet.ru

Received August 22, 2003

Abstract—An experimental study of the vertical acoustic sounding of stratified flows is carried out. The flows include two-dimensional wakes behind cylinders and three-dimensional wakes behind single ring vortices. A simultaneous optical visualization of the sound-scattering zones is performed. By the spectral processing of the schlieren patterns (in the wave number domain) and echo signal arrays (in the frequency domain), the most probable structure-forming elements of the scattering region are determined. These elements are internal boundary layers whose vertical size (thickness) is determined by the local frequency of internal waves and the kinematic viscosity. © *2005 Pleiades Publishing, Inc.*

Acoustic sounding is one of the most rapidly progressing methods for a fast evaluation of the physical properties and characteristics of the processes that occur in a stratified ocean or atmosphere [1]. Sounding is used to detect relatively large-scale (streams, flows, wakes behind obstacles, or internal waves) and also fine-structured phenomena. The interpretation of the results of measurements involves the problem of determining the nature of scatterers, which can be of a physical (bubbles, suspensions), biological, or hydromechanical origin (turbulence, multicomponent convection, vortices).

The identification of the class of scatterers is based on the empirical laws [1] established under full-scale and laboratory conditions. One can reliably detect the presence of gas bubbles (by the resonant frequencies in echo signals) and the presence of a biological suspension (from the frequency dependence and the characteristic temporal migration of the depth), while inorganic suspensions are less reliably identified.

Hydrodynamic mechanisms that lead to the formation of an acoustic contrast include thermohaline convection [2], turbulence, and more compact formations, such as vortex bundles, ring vortices, and coherent structures [3–5]. The scattering from turbulent wakes behind bodies was studied under half-full-scale conditions in [6]. Theoretical models of scattering in the ocean were developed using hypotheses that contain deterministic spectra of refraction-index fluctuations; such models can be extended to the case of a three-dimensional anisotropy of inhomogeneities [7].

Theoretical and experimental studies of the scattering from a stationary single two-dimensional vortex in a homogeneous fluid revealed zero levels in the forward and backward directions [5]. However, according to laboratory experiments in a stratified fluid, a compact vortex strongly scatters ultrasound in all directions [8]. In two-dimensional stratified wake flows, which may contain vortices, layered structures, and small-scale turbulence, peaks associated with specular reflection from high-gradient boundary surfaces arise in certain directions against the background of diffuse volume scattering [9]. It is of both theoretical and practical interest to study scattering from discrete vortex formations, whose variety underlies the classification of flow regimes behind obstacles.

From the viewpoint of the scattering of sound, the regions of interest in a fluid are those containing a set of elements with close characteristic dimensions. An example is a turbulence region, which, at a certain stage of evolution, is characterized by a statistically uniform internal scale.

In a stratified fluid, in addition to turbulence, other mechanisms generate small-scale inhomogeneities possessing such a property. One of them is the so-called internal boundary layer formed as a result of the interaction of internal waves with the fine structure of the flow.

The existence of internal boundary layers follows from the analysis of the system of equations of motion, which, for an isothermal case, can be reduced to the form [1]

$$\begin{aligned}
&\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\nabla)\mathbf{v} = -\frac{\nabla p}{\rho_0} + \mathbf{g} + \nu\nabla^2\mathbf{v},\\
&\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho\mathbf{v}) = 0,\\
&\frac{\partial s}{\partial t} + (\mathbf{v}\nabla)s = \kappa_S\nabla^2 s.
\end{aligned} \tag{1}$$

Here, $\mathbf{v}$ is the velocity of the fluid, p is the pressure, $\mathbf{g}$ is the acceleration of gravity, s is the salinity, ν is the kine-

matic viscosity, and κ_s is the salt diffusion coefficient. The boundary conditions are the no-slip and no-flux conditions at the surface of the moving object of height D (or a stationary object in a flow) and the decay of all disturbances at infinity.

A detailed analysis of system (1) shows that each type of internal waves corresponds to two different types of layers on surfaces in a flow [10] or on interlayers inside the fluid [11, 12], whose thicknesses are determined by the local frequency of internal waves and the corresponding kinetic coefficient, i.e., the kinematic viscosity or the salt diffusion coefficient.

The term "layer" means that this structural element has a certain vertical size (thickness), which, for a dynamic layer (i.e., a layer in the flow velocity field) is determined as [11]

$$\delta_\nu = \sqrt{2\nu/\omega_i},$$

where ω_i is the local frequency of internal waves. The horizontal length of the layer may take any value, which, in the general case, may be smaller than its thickness. The flow region, within which the viscosity and the frequency ω_i are approximately constant, should contain a set of elements characterized by the same vertical scale δ.

The efficiency of the backscattering of sound from a microstructured region depends on the relation between the structure-forming microscale δ (i.e., the internal scale, which is the predominant one) and the sound wavelength λ. As is known, maximum volume scattering corresponds to the condition $\delta = \lambda/2$ (the Bragg resonance), and, therefore, the closeness of the scale δ to the half-wavelength $\lambda/2$ can be accepted as the criterion for classifying a given inhomogeneity with the class of effective scatterers.

In this paper, with allowance for the aforementioned criterion, we study the role of internal boundary layers as sources of scattering on the basis of experimental data on the echo sounding and optical visualization of stratified flows.

We consider the scattering of sound characterized by frequency ω and wavelength λ from a hydrodynamic wake in an exponentially stratified fluid, whose density $\rho(z)$ linearly increases with depth due to the increase in salinity. The temperature is assumed to be constant. The stratification is characterized by the scale $\Lambda_\rho = |d(\ln\rho)/dz|^{-1}$ and the buoyancy frequency $N = \sqrt{g/\Lambda_\rho} = 2\pi/T_b$ (T_b is the buoyancy period), and the mechanical oscillations (internal waves) are characterized by the frequency ω_i and the wavelength λ_i.

The dimensionless parameters of the wake are as follows: the Reynolds number $\mathrm{Re} = D/\delta_u = UD/\nu$, where U is the velocity of motion of the wake-forming object (a cylinder or a vortex), and the internal Froude number $\mathrm{Fr} = U/ND$.

The simultaneous utilization of the schlieren visualization method [13] and acoustic sounding, which complement each other, extends the scope of the experiment. The optical image spectra allow for the determination of the spatial structural features of the flow and the structure-forming scales. The frequency spectra of oscillations of the echo signal make it possible to single out the local frequencies of internal waves, which, together with the kinetic coefficients, determine the current dimensions of the internal boundary layer in the insonified region.

The experiments were performed in a tank with the dimensions $240 \times 40 \times 60$ cm^3, which had transparent walls. The tank was filled with a linearly stratified water solution of common salt by the continuous displacement method. The density distribution profile and the buoyancy period were measured before each experiment by a conductivity sensor (with an error of no more than 5%). In the given series of experiments, the period was $T_b = 5.4$ s.

Two types of wake flows were generated in the tank: a two-dimensional wake behind a horizontally moving circular cylinder (with a diameter D of 1.5 or 5.0 cm) and a three-dimensional wake behind ring vortices. The cylinder was attached by thin knifes to a sledge, which moved along the guides mounted above the tank.

The vortices were generated by a spring–piston mechanism [8] with a horizontal guiding taper pipe ending with a nozzle. The vortex emerging from the nozzle moved with some velocity through the sonar operation zone, where it was insonified with acoustic signals. The axis of the schlieren visualization system and the axis of motion of the perturbing object (vortex or cylinder) were perpendicular to each other in the horizontal plane. The two axes intersected the vertical axis of the sound beam at the center of the tank. The error in the determination of the velocities of the cylinder and the vortex was no greater than 5%.

The flows were visualized by an IAB-458 schlieren visualization system in two ways: the conventional method (vertical slit–knife-edge combination) and the Maksutov method (slit–thread at the focus). With allowance for the linear relation between the density and the refraction index of the salt solution, the first method visualizes the horizontal component of the refraction index and the second method visualizes the magnitude of this component [13].

The buoyancy period was measured by the conductivity oscillations in the wake behind the density mark (no greater than 1 mm in thickness), which was formed by vertically rising gas bubbles or dropping sugar crystals. The error in measuring the buoyancy period was no greater than 10%.

The sonar antenna had the form of a circular piston radiator (a piezoceramic disk 2.5 cm in diameter),

which was tuned to the mechanical resonance frequency. The radiation power (in the pulse) was 0.3 W. The acoustic sounding was performed in the vertical direction by radio pulses with a duration of $\tau = 40$ μs at a wavelength of $\lambda = 0.15$ cm (a carrier frequency of $f_0 = 1$ MHz). The repetition period was chosen to be sufficiently large ($T = 0.16$ s), which allowed us to avoid a superposition of the multiple echo from the tank walls on the useful signal. The sampling interval in the digitization of the echo signal was $\Delta t = 2.2$ μs. The distance from the acoustic antenna plane to the cylinder center was $H = 23$ cm.

The schematic diagram of the acoustic sounding of a vortex ring and its wake is shown in Fig. 1, where the characteristic dimensions of the vortex ring and the sound signal are indicated.

The experiment was performed as follows. After the tank was filled and all inhomogeneities decayed, the buoyancy period was measured. Then, an inhomogeneity was generated by a towed cylinder or a moving vortex ring, and this inhomogeneity produced a contrast in both the acoustic and optical fields. The fully created wake flow was sounded with acoustic pulses and simultaneously photographed. The photograph was used together with the echo signal array in computer processing.

The acoustic data were used to determine the internal wave frequencies, which served to calculate the dimensions of the internal boundary layers. In the optical image processing, at the first stage, a rectangular fragment was separated with the dimensions Wh, where W is the mean transverse size of the sound beam and $h = c\tau/2$ is the length of the acoustic transmission. Then, for this fragment, a spectrum of vertical wave numbers was constructed, on the basis of which the presence of linear dimensions obtained from the acoustic data was verified and the relation between these dimensions and the wavelength was determined.

Figures 2a–2c show a sequence of schlieren patterns of the flow behind a circular cylinder with the diameter $D = 1.5$ cm. The images are obtained by the Maksutov method (the cylinder moves from right to left). The flow regime is the vortex-wave nonturbulent one, because the Reynolds number is much smaller than its critical value $\mathrm{Re}_c = 1000$. The central strip on the left-hand side of the frame (Fig. 2a) represents the blocked fluid at the depth where the body moves, and the inclined strips represent the crests (the dark line above) and troughs (the gray line below) of nonstationary internal waves, which smoothly transform to the attached internal waves behind the body.

With the given parameters of motion (Fr = 1.6 and Re = 410; see Fig. 2a), a vortex density wake is formed behind the cylinder. As the vertical velocity component decays, the three-dimensional vortices transform to a plane-layered structure. In this case, the layer with the highest optical contrast lies in the plane of motion of the cylinder center. The height of the wake and the number of interlayers in it increase with time.

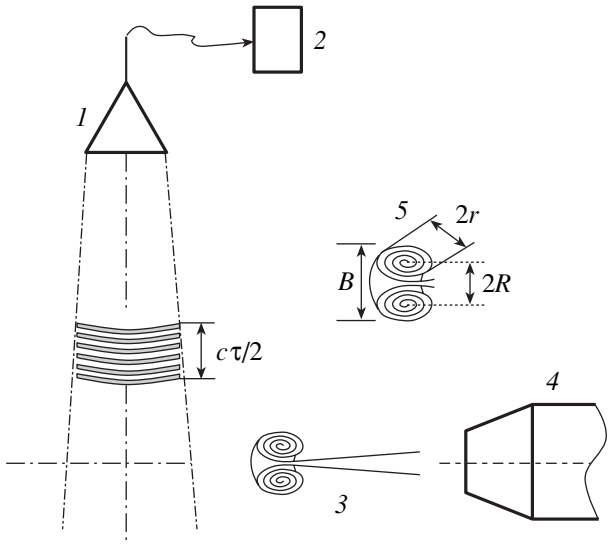


Fig. 1. Schematic diagram of the experiments on the acoustic sounding of a vortex ring and its wake: (*1*) sonar antenna, (*2*) computerized sonar, (*3*) vortex ring emerging from the nozzle, and (*4*, *5*) geometric parameters of the vortex ring.

To make the processing and the representation of the results more convenient, the echo signals are combined into a two-dimensional array, whose elements A_{mn} are proportional to the scattering amplitude and the numbers of rows m and columns n of which are proportional to the depth $z_m = z_0 + (m - 1)c\tau/2$ and the running time $t_n = (n - 1)T$. The array is visualized in the form of an echogram (Fig. 2d) in the dimensionless time-versus-depth coordinates. The amplitude of the echo signal is represented by the degree of blackening.

From the schlieren images and the echogram, one can see that the density wake and individual high-gradient interlayers perform oscillations in depth. To extract these oscillations from each column of the array, we select the maximum value of A_{mn} with its pair of coordinates z_m and t_n, after which we filter out the pairs with $A_{mn} < A_c$, where A_c is the threshold value.

The resulting array (z_m, t_n) represents the time dependence of the vertical coordinate of the acoustic mark with maximum contrast (Fig. 3a). Its frequency spectrum contains three maxima in the internal wave range ($f^* < 1$), which correspond to the frequencies $\omega_i = 0.15$, 0.53, and 1.16 s^{-1} indicated in increasing order in Fig. 3b.

For these frequencies, the thickness of the dynamic internal boundary layers is equal to $\delta_\nu = \sqrt{2\nu/\omega_i} = 0.36$, 0.19, and 0.13 cm, respectively. The latter size is close to the half-wavelength of sound, and, therefore, it is the most probable structure-forming element of the

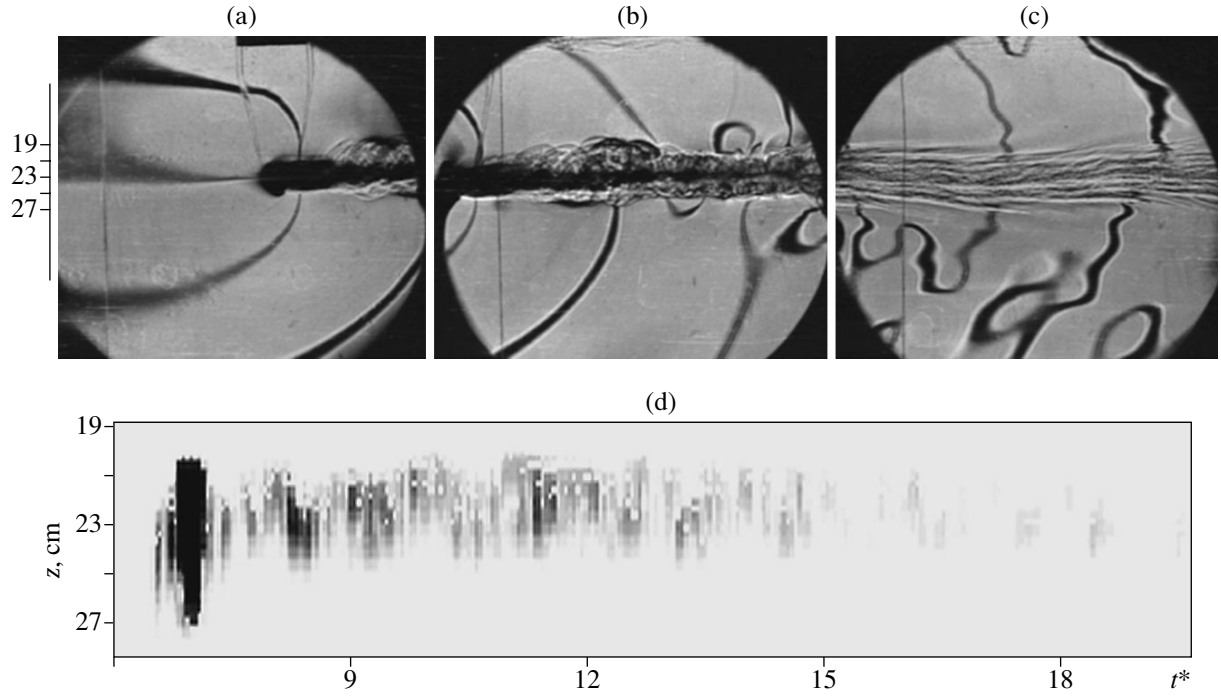


Fig. 2. Schlieren pattern (side view) of the flow behind a cylinder for t^* = (a) 6.5, (b) 8.0, and (c) 11.0, (a–c) T_b = 5.4 s, D = 1.5 cm, U = 2.8 cm/s, Fr = 1.6, and Re = 410. (d) The corresponding scattering echogram as a function of the dimensionless age t^*. The elongated light-striking near t^* = 7 is the echo from the cylinder, which is formed by the main and side lobes of the directivity pattern of the sonar antenna.

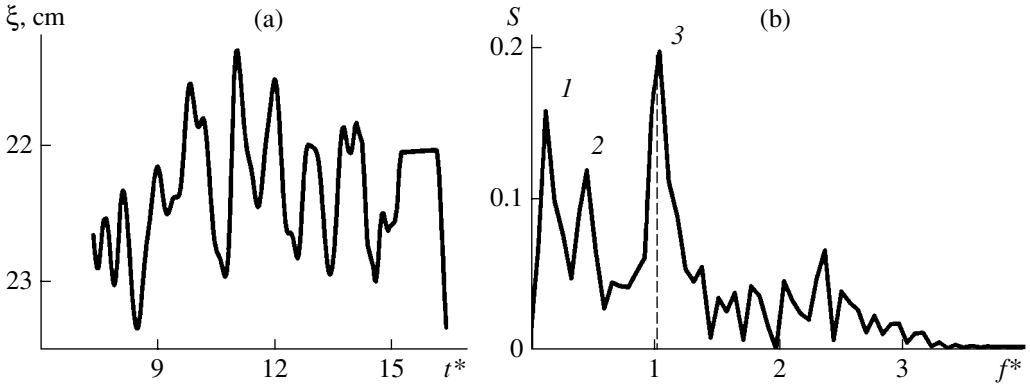


Fig. 3. (a) Record of oscillations and (b) the frequency spectrum of the depth at which the central sound-scattering layer lies in the wake behind the cylinder: T_b = 5.4 s, D = 1.5 cm, U = 2.8 cm/s, Fr = 1.6, and Re = 410 (the dimensionless time and frequency are $t^* = t/T_b$ and $f^* = fT_b$).

scattering region. This is also evidenced by the fact that the height of spectral peak *3* observed at the frequency of 1.11 s^{-1} (Fig. 3a) is maximal. In the acoustic sounding of stratified flows, the backscattering frequency spectra often exhibit peaks at the double frequency of internal waves as well, which is explained by the insensitivity of the scattering amplitude to the sign of the scattering acoustic impedance jump [14].

When the diameter of the cylinder considerably increases, the wake region expands, and contrast regions are formed on both sides of the axis of motion at a noticeable distance from it (Figs. 4a–4d). The wave motions in the upper and lower halves of the wake are in antiphase, which manifests itself in the oscillations of the acoustic contrast (Fig. 4e). The frequency spectra of the upper and lower halves differ little in shape (Fig. 4f), and, in the internal wave frequency range ($f^* < 1$), each of them has a principal maximum at the frequency $\omega_i = 2\pi \times 0.6/T_b = 0.5$ s^{-1}. However, the width of the maxima is much greater than in the first experiment (Fig. 2). This is related to the greater size of the wake in the vertical and, hence, a noticeable vari-

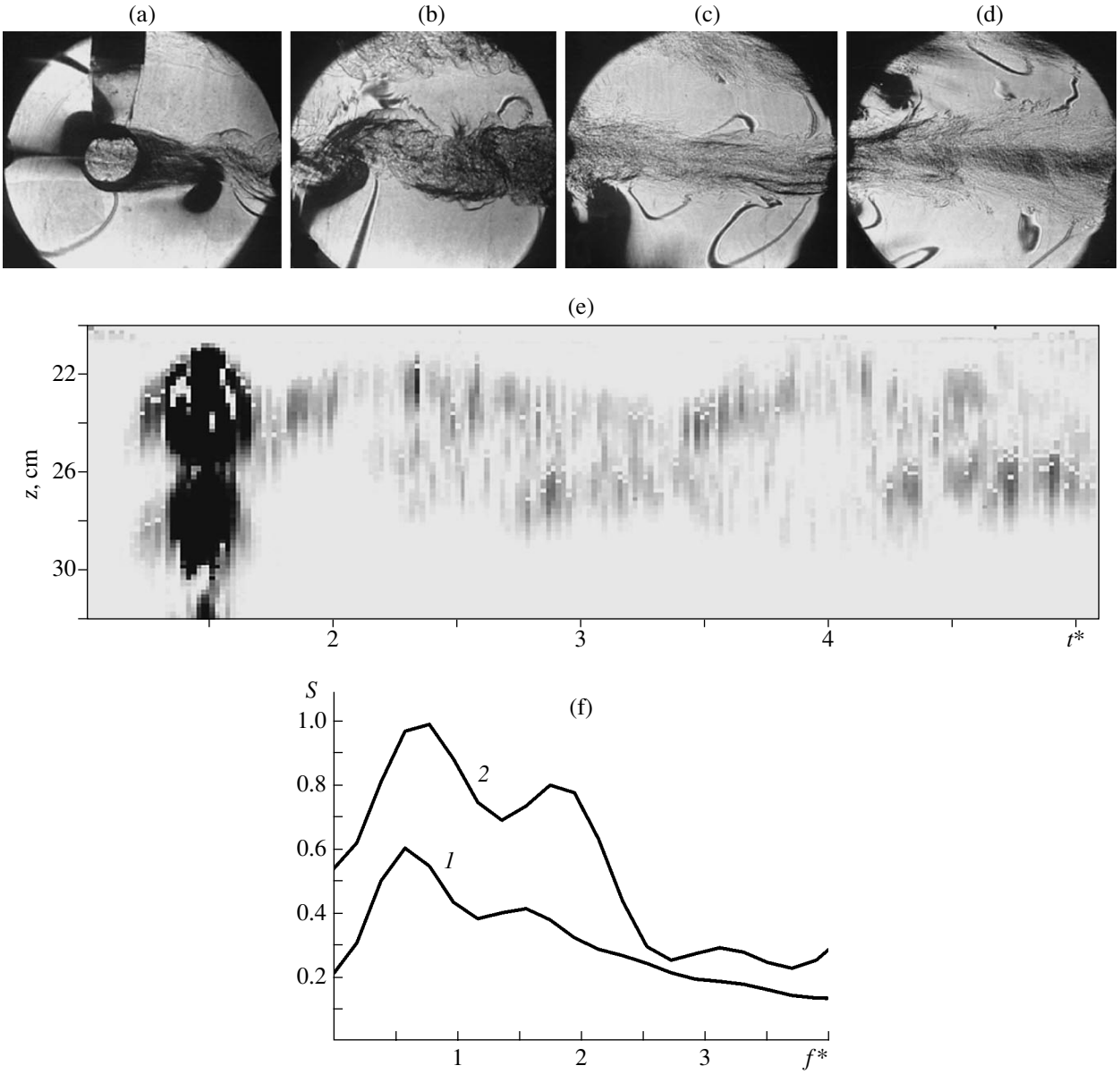


Fig. 4. Schlieren pattern for the wake age $t^* = (a)$ 1.7, (b) 2.2, (c) 3.5, and (d) 4.0; (e) the echogram; and (f) the frequency spectrum of the echo signal in the (*1*) upper ($z < 24$ cm) and (*2*) lower part ($z > 24$ cm) of the wake behind a cylinder: $D = 5$ cm, $T_b = 7.8$ s, $U = 3.6$ cm/s, Re = 1900, and Fr = 0.9. The elongated light-striking observed in the echogram with the center at $t^* = 1.5$ is the echo from the moving cylinder.

ability of the internal wave frequencies from point to point, according to the dispersion relation $\omega_i = N\cos\theta$, where θ is the slope angle of the phase surface of the wave with respect to the vertical. The values of the frequency ω_i (at a level of 0.7 of the first maxima of curves *1* and *2*, Fig. 4f) lie within 0.3–0.7 s^{-1}, and, according to Eq. (2), the ranges of the dimensions of the concentration (salt) and dynamic layers should be $\delta_s = \sqrt{2\kappa_s/\omega_i} = 0.005$–$0.01$ and $\delta_\nu = \sqrt{2\nu_s/\omega_i} = 0.16$–$0.25$ cm, respectively.

The actual presence of structural elements with these dimensions in the wake can be judged from the spatial spectra of schlieren patterns, which are constructed for the vertical wave number k_z:

$$F_z(k_z, n) = F_z(k_z, x) = \sum_{m=1}^{m=M} A_{mn}\exp[ik_z m\Delta z]\Delta z, \quad (3)$$

where $x = U(n-1)T$ is the horizontal separation of the point of scattering from the cylinder center, $\Delta z = c\Delta t/2$, and $M = \mathrm{int}(\tau/\Delta t)$ is the number of rows that fit within

the sounding pulse length $c\tau/2$. Function (3) represents the dependence of one-dimensional spectra for each column vector of the scattering matrix on the local coordinate x. To reduce this dependence to the final form, taking into account the inhomogeneity of the sound beam, it should be averaged over the horizontal coordinate with an interval equal to the width of the Fresnel zone $\mu = \sqrt{H\lambda} \approx 2$ cm:

$$Y_z(k_z, x) = \frac{1}{n_0} \sum_{n}^{n+n_0-1} F_z(k_z, n), \qquad (4)$$

where $n_0 = \mathrm{int}(\mu/UT)$.

Spectral form (4) allows us to take into account the influence of the horizontal length of the scale inhomogeneities $l_z = k_z^{-1}$ and the width of the sound beam on the echo formation.

For better illustration, Figs. 5a and 5d show magnified fragments of the flow patterns at the stages $t^* = 3.5$ and 4.0. Here, on the background of small-scale inhomogeneities, inclined high-gradient interlayers extending along the horizontal are observed. As the flow becomes degenerate, the stratification is recovered within the whole thickness of the fluid and internal waves begin to penetrate into the wake. In the photograph, this manifests itself as alternating dark spots along the central horizontal in Fig. 5d.

Averaged spectra (4) of optical inhomogeneities can be conveniently represented as a grayscale image in the (x, k_z) coordinate plane. The degree of blackening in Figs. 5b and 5e is proportional to the scattering level A_{mn}.

The significant spectral amplitudes exceeding the threshold value of 0.3 lie in the region of wave numbers $k_z < 8$ cm^{-1}. Thus, the region of existence of the salt layer ($k_z = 100$–200 cm^{-1}), as well as the resonance scale $\lambda/2 = 0.075$ cm ($k_z = 13.4$ cm^{-1}), are far beyond the limits of the significant part of the spectrum.

At the same time, the range with the lower boundary $k_z = 8$ cm^{-1} completely covers the limits of variability of the dynamic layer ($k_z = 4$–6.25 cm^{-1}). The most probable scatterer is the structure with the internal boundary layer scale of 0.16 cm, which is closest to the size of $\lambda/2$.

Since oceanic currents are predominantly three-dimensional, a question arises as to the presence of internal boundary layers in three-dimensional structures and their participation in the scattering of sound. Such a structure may be a moving vortex ring, which is an effective sound scatterer by itself and which also produces a density wake with a high acoustic contrast [8].

The wake formed behind a vortex ring contains the same elements as those in the wake behind an obstacle: the attached internal waves, the advance disturbances including the zero-frequency waves, and the density wake. The dimensionless parameters characterizing the regime of motion are the Reynolds and Froude numbers determined by the external vertical size B (Fig. 1) and the initial translational velocity U of the vortex: $\mathrm{Re}_e = UB/\nu$ and $\mathrm{Fr}_e = U/NB$. An important internal parameter that determines the stability of the vortex as a whole is the ratio η of the orbital velocity u_φ to the translational velocity U. Taking into account the weakness of the stratification, the value of this ratio can be estimated by the relation suitable for a homogeneous fluid [15]:

$$\eta = \frac{u_\varphi}{u} = \frac{2\beta}{\log(8\beta - 0.25)}, \qquad (5)$$

where $\beta = R/r$ is the ratio of the large and small radii of the vortex (Fig. 1).

The schlieren images of two different vortex rings with different parameters of motion are shown in Fig. 6. Sequential photographs presented in Figs. 6a, 6b, 6d, and 6e were obtained at an interval of 2 s.

In the first pair of images (Figs. 6a, 6b), one can see a structurally ordered compact vortex ring, which is shown in Fig. 6c on an enlarged scale. The second ring, characterized by a smaller value of the Reynolds number, forms a wake with a greater vertical size as compared to the first ring. The wake from the first ring is more ordered and has a contour formed by inclined high-gradient interlayers. Their shape testifies that the point of separation of the interlayer from the wake contour is uniformly displaced toward the axis of motion.

The wake of the second vortex consists of a sequence of ring structures formed by a set of small vortex rings connected by thin bundles. Its vertical size exceeds the height of the vortex itself (Figs. 6d, 6e). At the center of the upper half of Fig. 6e, one can see a light vertical line, which is the density wake of a rising air bubble.

Finer details are visible in the magnified fragments (Figs. 6c, 6f), where internal rings can be distinguished through the vortex shell (these rings are indicated by dashes). From these data, we calculated the mean values of the radii R and r and the ratios β and η. The resulting values are shown in the table. The ratio η for the first vortex is much greater than for the second one, which explains its stability. In the second vortex with a smaller value of η, a fast erosion of the shell of the internal surface takes place. As a result, an increase occurs in the amount of liquid transferred by the vortex to the wake in the form of vortex bunches.

The geometric difference between the flows manifests itself in the backscattering field (Fig. 7). In the first case, the scattering of sound from the vortex alone is detected. In the second case, an acoustic contrast is observed from both the vortex and its wake (a series of dark spots along the time axis) within a period of $t \approx 4T_b$. In Fig. 7b, the most pronounced feature is the dark inclined strip, which represents the acoustic illumina-

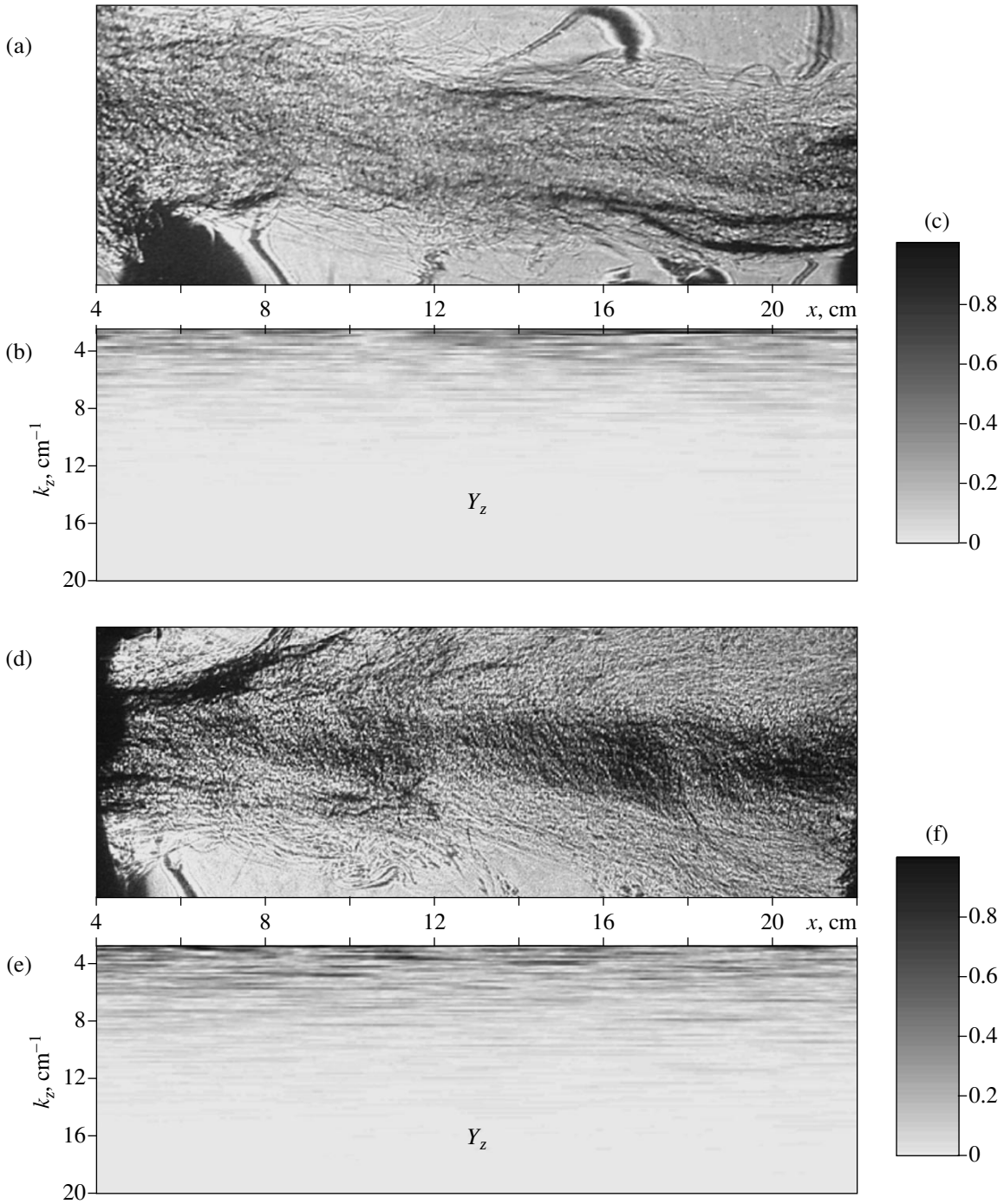


Fig. 5. Fragments of the schlieren patterns of the wake for t^* = (a) 3.5 and (d) 4.0. (b, e) The corresponding gray scale images of the amplitude spectra as functions of the vertical wave number k_z and horizontal coordinate x. (c, f) Spectral amplitude scales.

tion of a rising bubble; its angle of inclination determines the velocity of the bubble rise: 3.2 cm/s.

The comparison of the schlieren patterns behind the vortices (Fig. 6) testifies that the origin of the drastic difference in the acoustic patterns is the difference in the structural elements participating in the scattering. The wake in Fig. 6e is 3 cm in height, which coincides with the length of the echo sounding pulse $c\tau/2$ and by more than two times exceeds the vertical size of the other wake (Fig. 6b). In addition, from the schlieren images, we can conclude that the concentration of microscale inhomogeneities is much greater in the second wake. Thus, the number of scatterers in the region of intersection of the second wake with the sound pulse is sufficient for a noticeable acoustic contrast (Fig. 7b),

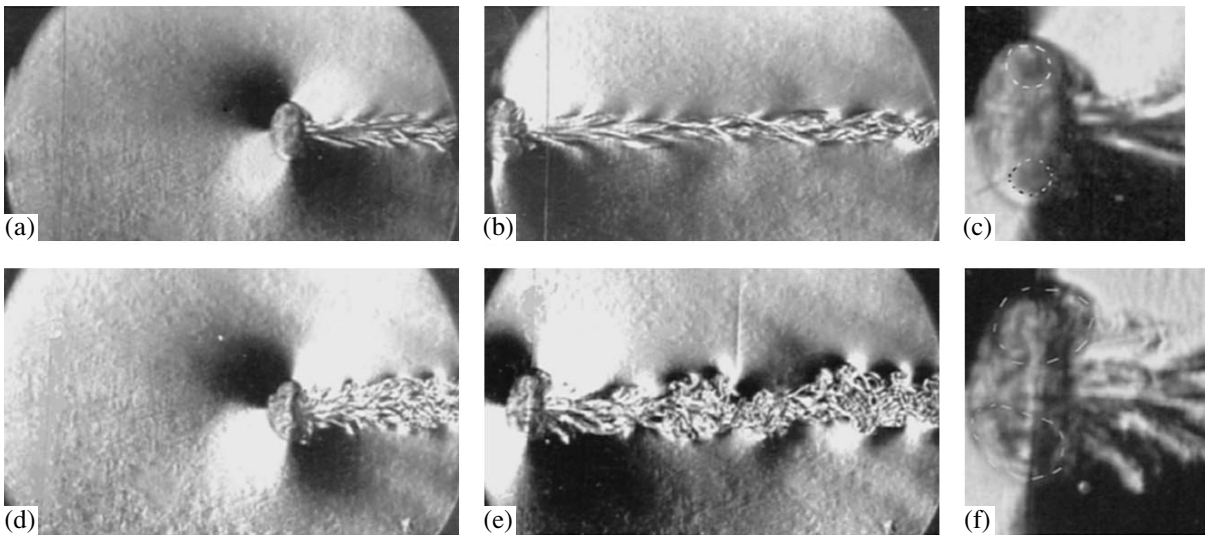


Fig. 6. (a, b, d, e) Schlieren patterns of the wake behind the ring vortices and (c, f) the magnified images of the vortices with dashes indicating the contours of the internal rings. The parameters of motion are T_b = 6.7 s, U = (a–c) 9.5 and (d–f) 8.5 cm/s, B = (a–f) 3 cm, Re_e = (a–c) 2800 and (d–f) 2500, and Fr_e = (a–c) 3.5 and (d–f) 3.1.

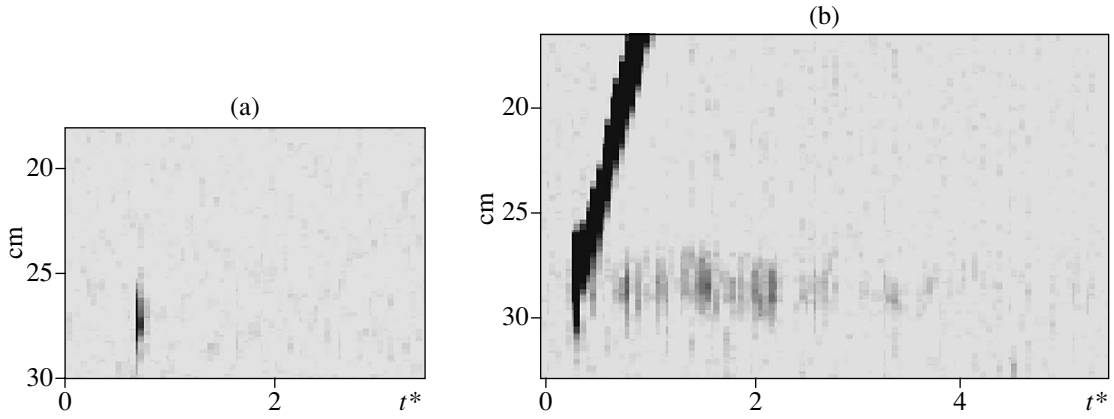


Fig. 7. Echograms of the wake behind the ring vortices for the flow regimes with Re_e = (a) 2800 and (b) 2500 and Fr_e = (a) 3.5 and (b) 3.1. The broad inclined strip in the echogram represents the scattering from a rising bubble.

while, in the first case, their number is too small to exceed the acoustic background level.

To reveal the structural inhomogeneities participating in the scattering, the same method as that used for two-dimensional flows is suitable, namely, a spectral analysis in the form of Eqs. (3) and (4) with a spectrum representation on the gray scale. However, in the case of a three-dimensional wake, it is necessary to take into account that the schlieren visualization system gives the projection of all inhomogeneities onto the image plane. This does not prevent the detection of fine vertical scales, but their distribution along the axis perpendicular to the image plane remains unknown.

With allowance for the above consideration, the spectra of the vertical wave number of a three-dimensional flow should be interpreted as indicators of the presence or absence of inhomogeneities with preset scales. From the form of the spectral fields (Fig. 8), it follows that both wake flows contain components with the dimensions of the order of the sound wavelength. Their number and type of distribution in the horizontal are approximately the same in both cases. Here, the wave number of 7 cm^{-1} (the scale of 0.14 cm, which is closest to the half-wavelength) is in fact the boundary, above which the spectral amplitudes are negligibly small.

Table

Plot in Fig. 6	R, cm	r, cm	β	η
c	1	0.35	2.8	4.2
f	1	0.65	1.5	2.8

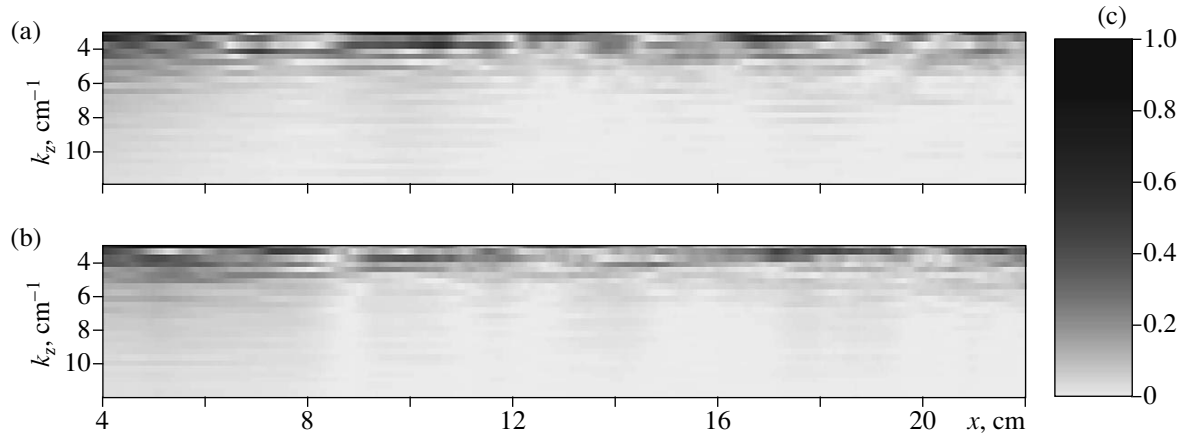


Fig. 8. Amplitude spectra as functions of the vertical wave number k_z and horizontal coordinate x. The spectra are calculated from the schlieren pattern of the wake behind a vortex ring: Re = (a) 2800 and (b) 2500 and η = (a) 4.2 and (b) 2.8. (c) The spectral amplitude scale.

The thickness of the internal boundary layer, when calculated from the local frequency of internal waves, is determined, as in the case of two-dimensional flows, according to Eq. (2) and the frequency spectrum of oscillations of the highest contrast acoustic mark (Fig. 9). In the case under consideration, we have three peaks, only one of which falls within the range of existence of internal waves: $\omega_i = 0.62$ s^{-1} ($f^* = 0.66$). The next peak, which originates from the insensitivity of the echo power to the sign of the density jump, occurs at the double frequency of the fundamental internal oscillation of the wake ω_i ($f^* = 1.32$, Fig. 9).

The third peak has a considerable amplitude, which is second in magnitude, but this peak is of no wave nature, because its dimensionless frequency $fT_b = 2.33$ lies far beyond the limits of the wave range. One of the mechanisms of the echo formation at such frequencies is the horizontal motion of a large-scale periodic structure through the sound zone [14]. In Fig. 6e, nodes and antinodes are seen in the wake with their average period L_c being equal to 3 cm. This yields the velocity of the wake motion $U_c = fL_c = f^*L_c/T_b = 1$ cm/s, which is much smaller than the average velocity of motion of the vortex itself (8.5 cm/s).

Taking the first value $\omega_i = 0.62$ s^{-1} as the fundamental frequency of internal waves, we obtain the thickness of the dynamic layer $\delta_\nu = 0.18$ cm. Of all possible dimensions, this one is closest to $\lambda/2$ and, in addition, it falls within the range of the actual scales of spatial spectra of the schlieren pattern of the wake (Fig. 8). These features testify that the basis of the structural inhomogeneities responsible for the sound scattering is formed by the dynamic internal boundary layers.

Thus, in the experiments on the backscattering of ultrasound from wake flows behind a horizontal cylinder and a single vortex ring, the general regularities of the acoustic contrast formation are determined.

The analysis of the experiments shows that the most probable sources of scattering are the internal boundary layers in the velocity field, with the scale of these layers determined as $\sqrt{2\nu/\omega_i}$.

The extension of the laboratory data to full-scale conditions is possible with supporting the criterion of modeling, namely, the ratio of the thickness of the internal boundary layer to the sound wavelength, which in the experiments proved to be close to unity. Taking the quantity $\sqrt{2\nu/N}$ as the estimate of the minimum thickness of an internal boundary layer, we obtain that, under typical oceanic conditions ($\nu = 0.01$ cm^2 s^{-1}, $N = 5 \times 10^{-3}$ s^{-1}), the middle of the frequency range within

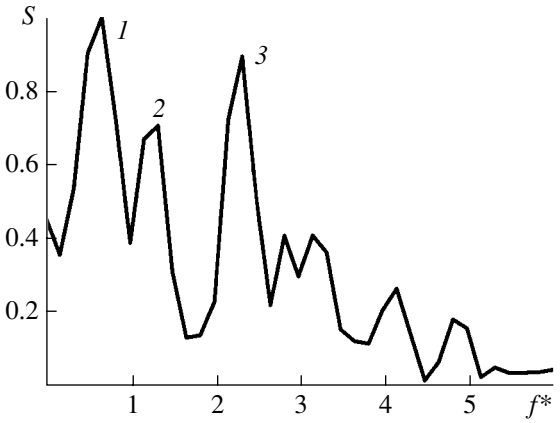


Fig. 9. Frequency spectrum of the vertical oscillations of the acoustic mark in the wake behind the vortex ring; Re = 2500. The maxima in the curve correspond to (*1*) the frequency of internal waves, (*2*) the double frequency of internal waves, and (*3*) the frequency of the echo signal oscillations due to the motion of the periodic structure of the wake through the sound zone.

which the most intense backscattering from internal boundary layers is expected lies in the region of 75 kHz.

ACKNOWLEDGMENTS

This work was supported by the Russian Academy of Sciences (the Fundamental Research Program "Dynamics and Acoustics of Inhomogeneous Fluids, Gas–Fluid Systems, and Suspensions" financially supported by the Division of Power, Engineering, Mechanics, and Control Processes of the Russian Academy of Sciences), the Ministry of Industry, Science, and Technology of the Russian Federation (the Program in Support of Unique Facilities), the Ministry of Education of the Russian Federation (the Federal Special-Purpose Program "Integration," project no. 02-Ya0058/993), and the Russian Foundation for Basic Research (project no. 03-01-00434).

REFERENCES

1. *Physics of the Ocean*, Vol. 2: *Hydrodynamics of the Ocean* (Nauka, Moscow, 1978) [in Russian].
2. V. P. Shevtsov and A. S. Salomatin, Okeanologiya (Moscow) **32** (4), 661 (1992).
3. M. H. Orr, L. R. Haury, P. H. Wiebe, and M. G. Briscoe, J. Acoust. Soc. Am. **108**, 1595 (2000).
4. V. E. Ostashev, T. M. Georges, and G. H. Goedecke, J. Acoust. Soc. Am. **109**, 2682 (2001).
5. T. Colonius, S. A. Lele, and P. Moin, J. Fluid Mech. **260**, 271 (1994).
6. I. Pelech, G. Zipfel, and R. Holford, J. Acoust. Soc. Am. **73**, 528 (1983).
7. E. A. Kopyl, Yu. P. Lysanov, and L. M. Lyamshev, Akust. Zh. **48**, 517 (2002) [Acoust. Phys. **48**, 453 (2002)].
8. V. V. Mitkin, V. E. Prokhorov, and Yu. D. Chashechkin, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 6, 92 (2001).
9. V. V. Mitkin, V. E. Prokhorov, and Yu. D. Chashechkin, Dokl. Akad. Nauk **377** (2), 201 (2001) [Dokl. Phys. **46**, 202 (2001)].
10. A. V. Kistovich and Yu. D. Chashechkin, Prikl. Mat. Mekh. **57** (4), 50 (1993).
11. Yu. V. Kistovich and Yu. D. Chashechkin, Prikl. Mat. Mekh. **39** (5), 88 (1998).
12. Yu. V. Kistovich and Yu. D. Chashechkin, Preprint No. 674, IPM RAN (Inst. for Problems in Mechanics, Russian Academy of Sciences, 2001).
13. Yu. D. Chashechkin, J. Visual. **1** (4), 345 (1999).
14. V. E. Prokhorov and Yu. D. Chashechkin, Akust. Zh. **41**, 908 (1995) [Acoust. Phys. **41**, 804 (1995)].
15. H. Lamb, *Hydrodynamics*, 6th ed. (Cambridge Univ. Press, Cambridge, 1932; Gostekhizdat, Moscow, 1947).

Translated by E. Golyamina