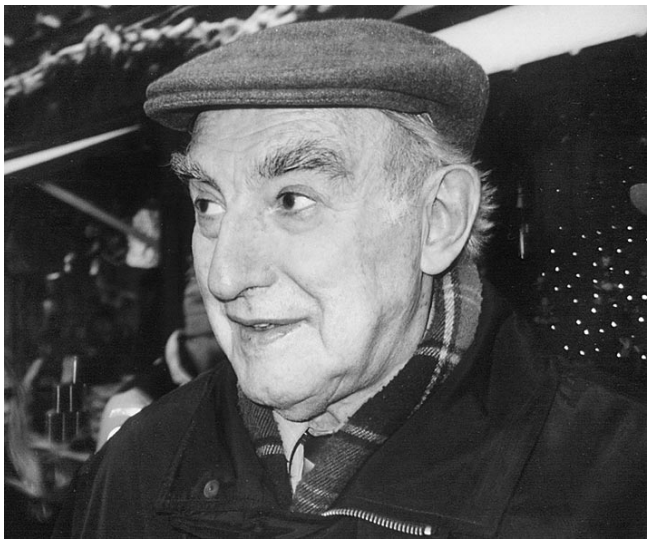
Acoustical Physics, Vol. 50, No. 3, 2004, pp. 239–242. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 293–297.

Vitaliĭ Lazarevich Ginzburg—Winner of the Nobel Prize



At the end of 2003, an event of historical significance for all of Russian Science has occurred. Academician V.L. Ginzburg—an outstanding scientist and one of the most respected physicists—was awarded the 2003 Nobel Prize in physics "For pioneering contributions to the theory of superconductors and superfluids." Ginzburg shared the prize with his younger colleagues A.A. Abrikosov and J. Leggett.

The editorial board of *Acoustical Physics* celebrates this event not only for formal reasons (because acoustics is part of physics) but also in connection with the fact that Ginzburg earlier carried out long-term studies concerned with the propagation of electromagnetic and acoustic waves. One should remember his monographs *Theory of Radio Wave Propagation in Ionosphere* (GITTL, Moscow, 1949) and *Propagation of Radio Waves* (co-authored with Ya.L. Al'pert and E.L. Feinberg; GITTL, Moscow, 1953), as well as his paper "On the General Relation between the Absorption and Dispersion of Sound Waves," which was published in 1955 in the first issue of the newly founded journal *Akusticheskii Zhurnal* (Akust. Zh. **1**, 31 (1955) [Sov. Phys. Acoust. **1**, 32 (1955)]).

The author of the present article (O.V. Rudenko) was lucky enough to attend the Nobel Prize Ceremony in Stockholm, where Ginzburg received his award. This occasion made it possible to obtain some original materials which may animate the official tribute.

Studies related to the detection and investigation of macroscopic quantum phenomena at low temperatures had been acknowledged by the Nobel Committee as great scientific achievements years ago. In 1913, the Nobel Prize was given to H. Kamerlingh Onnes "for investigations on the properties of matter at low temperatures which led, *inter alia*, to the production of liquid helium." In 1962, the prize was given to the Soviet physicist L.D. Landau "for pioneering theories for condensed matter, especially liquid helium." In 1972, J. Bardeen, L. Cooper, and J. Schrieffer were awarded the Nobel Prize "for their theory of superconductivity, usually called the BCS-theory." In 1978, the prize was given to another Soviet physicist P.L. Kapitsa "for his basic inventions and discoveries in the area of low-temperature physics."

The works by V.L. Ginzburg on low-temperature physics had been widely known for more than 50 years. In his paper "On the theory of Superconductivity" (co-authored with L.D. Landau; Zh. Eksp. Teor. Fiz. **20**, 1064 (1950)) the order parameter used for describing the phase transition to the superconducting state is taken in the form of a complex function, which plays the role of a certain effective wave function of superconducting electrons. In the same paper, the well-known Ginzburg–Landau nonlinear equation is proposed. This equation is of fundamental significance for many fields of condensed matter physics and astrophysics. However, it may be of interest to know what exactly were the main arguments that governed the decision of the Nobel Committee in Physics to give the prize to Ginzburg, Abrikosov, and Leggett.

At the Prize Award Ceremony held on December 10, 2003, the chairman of the Nobel Committee, Professor

1063-7710/04/5003-0239$26.00

Mats Jonson, said that this prize was entirely devoted to order that occurs not in the common surrounding world but in the microworld of electrons and atoms. This world is governed by the laws of quantum physics and exhibits many unusual phenomena unobservable in the macroworld. However, in some cases, a special type of order among electrons and atoms may enhance the quantum phenomena of the microworld to make them visible to the unaided eye. The 2003 Nobel Prize was given to the scientists whose works were of crucial importance for the understanding of the relations between order and macroscopic quantum phenomena called superconductivity and superfluidity.

As early as 1911, the Dutch physicist Kamerlingh Onnes found that the electrical resistance of mercury disappears at a temperature of several degrees above absolute zero. He called this phenomenon superconductivity. Later, Russian physicist Petr Kapitsa introduced a similar term "superfluidity" to describe the flow of liquid helium at even lower temperatures, when the flow occurred without friction or resistance.

From the very beginning, it was evident that superconductivity could find wide application in modern industry. For example, by making coils of superconducting wires, one could make high-power electromagnets operating without an energy loss. Unfortunately, the first superconductors returned to the state of normal metal even in weak magnetic fields. Later, superconductors of other types were discovered. They allowed a coexistence of superconductivity and magnetism and retained superconducting properties even in high magnetic fields. Superconducting magnets designed owing to these findings have found application in high-resolution magnetic imaging devices for medical diagnostics and in the equipment of large accelerators in elementary particle physics.

Two winners of the Nobel Prize, Ginzburg ad Abrikosov, made a decisive contribution to the understanding of the nature of coexistence of superconductivity and magnetism. Ginzburg, together with Landau, developed a theory that described in detail how superconductivity disappears when certain critical values of electric current and magnetic field are reached. Ginzburg and Landau introduced a measure of electron order and called it the order parameter. Governed by physical intuition, they formulated mathematical equations whose solution determines the order in a superconductor. They achieved a good agreement with the results of experiments on superconductors available at that time. The Ginzburg–Landau theory proved to be rather general and widely applicable, and it is still used today in many fields of physics.

Soon, experiments with new superconducting materials were carried out, and these experiments provided unexpected results. It was Abrikosov who discovered more complex types of order, which account for the aforementioned deviations. By performing a comprehensive analysis of Ginzburg–Landau equations, he showed how the spatial distribution of the order parameter is formed by vortices and how magnetic field can penetrate into a superconductor owing to these vortices. He explained the mechanism of coexistence of superconductivity with magnetism in a new class of materials called type-II superconductors. This was breakthrough in the studies of new superconducting materials.

An order of an even more complex type takes place in superfluid helium III, which was discovered in the early 1970s. This occurred long after Kapitsa's discovery of superfluidity in helium of a natural origin, in which the heavy helium IV isotope predominates. The two helium isotopes are fundamentally different. Helium IV belongs to the class of particles called bosons, which can be directly ordered to form a superfluid through the so-called Bose–Einstein condensation. By contrast, helium III is a fermion, like an electron. Such particles should form pairs before reaching the superfluid state. It was found that, in liquid helium III, the atoms form pairs with internal degrees of freedom. This occurs partially because of the rotation of the atoms with respect to each other and partially because of their magnetic properties or their spin. As a consequence, in this case, the Ginzburg–Landau order parameter has eighteen components rather than two (as in superconductors), and the resulting superfluid is anisotropic: it may have different properties in different directions.

It was Leggett who managed to explain the relation between the properties of the new superfluids and the variety of the types of order allowed by the multicomponent order parameter. His theory made it possible to interpret the results of experiments with helium III. It was also used in other fields of science, in various problems of the physics of liquid crystals, and in cosmology.

These were the arguments considered by the Nobel Committee when making the decision to award Ginzburg, Abrikosov, and Leggett the 2003 Nobel Prize in physics for their pioneering contributions to the theory of superconductors and superfluids.

The Prize Award Ceremony was the culmination of Nobel Week, which lasted from December 6 to December 13 and included many speeches, meetings, and receptions. The most important events were as follows: a press conference of the Nobel Prize winners at the Royal Swedish Academy of Sciences (December 7), Nobel lecture at Stockholm University (December 8), TV discussion with the Nobel Prize winners and the reception by the Russian Ambassador to Sweden (December 9), Prize Award Ceremony at the Stockholm Concert Hall and the Nobel banquet at the City Hall (December 10), reception by the King and Queen at the royal palace (December 11), lecture and reception at the Royal Institute of Technology (December 12), a visit with a lecture to the ancient Uppsala University of Sweden and a dinner at Stockholm University in honor of St. Lucia (December 13). The main events were

widely reported in the media. The Nobel Prize Ceremony was watched on TV in real time by about a billion people (a quantity close to that for the final game of a world football cup). Unfortunately, the Russian TV channels were more formal towards the triumph of Russian science: they only reported on the subject in the news.

Briefly, the biography of V.L. Ginzburg is as follows. He was born in 1916 in Moscow. In 1938, he graduated from the Faculty of Physics of Moscow State University. In 1940, he received the degree of candidate of science, and two years later, the degree of doctor of science. Ginzburg was elected a corresponding member of the Academy of Sciences of the USSR in 1953, and in 1966 he became a full member of the Academy. Since 1940, he has worked at the Lebedev Physical Institute of the Academy of Sciences. In addition, from 1945 to 1968, he worked at the Lobachevski University, Gorki (Nizhni Novgorod), where he organized the Department of Radio Wave Propagation at the Radiophysical Faculty. He headed this department from 1945 to 1962. Since 1968, he has chaired the Department of Problems of Physics and Astrophysics at the Moscow Institute of Physics and Technology. Since 1998, he has also been Editor-in-chief of *Uspekhi Fizicheskikh Nauk* (*Physics–Uspekhi*).

Ginzburg received a State Award in 1953 and a Lenin Prize in 1966. He also received Mandel'shtam and Lomonosov academic awards in 1947 and 1962, respectively. He was awarded a Vavilov Gold Medal and a Lomonosov Large Gold Medal (1995), as well as many other orders and medals from the government. He was elected a member of the American National Academy of Sciences in 1981 and member of the Royal Society of London in 1987. He is a foreign member of many other academies and societies and has received many foreign awards for his scientific achievements.

Personally, Ginzburg wrote: "All this is rather conditional. Even worthless individuals may have formal achievements. However, scientific results are different, they are objective. … in superconductivity, superfluidity, ferroelectricity, the Vavilov–Cherenkov effect and transient radiation, radio astronomy, the origin of cosmic rays, scattering of light, and some other fields, my contributions are significant." This and subsequent citations are taken from the book by Ginzburg *On Science, Myself, and Others* (Fizmatlit, Moscow, 2003).

One of Ginzburg's papers included in this book ("Reminiscences of My Participation in the Atomic Project") is devoted to an almost unknown but important fact of his biography. In 1947, I.V. Kurchatov (the supervisor of the Atomic project) invited I.E. Tamm

Ginzburg receives the Nobel Prize (December 10, 2003).

(the future winner of the 1958 Nobel Prize for the theory of the Vavilov–Cherenkov effect) to participate in the project. "At that time, I was a five-year-old doctor of science and the deputy head of the Theoretical Department of the Lebedev Physical Institute while Tamm was the head of this department. Naturally, Tamm asked me to take part in the project … It was then that the thought occurred to me to use the reaction $^6\mathrm{Li} + n \longrightarrow t + {}^4\mathrm{He} + 4.6$ MeV. After many years, B.P. Konstantinov told me that he headed the construction of the plant for the separation of ^{6}Li… When on August 12, 1953, the first Soviet hydrogen bomb containing ^{6}Li was exploded, Americans detected ^{6}Li in the atmosphere and were amazed—this I read somewhere. Thus, *post factum*, many years after 1953, I understood that my proposal played an important part in our Atomic project, while my closer participation in this project was not allowed. … Presumably, I was not sent to the object because I was politically suspect. … After this, I never took part in any secret works but was considered as "secret" until 1987, i.e., for 32 years … In this connection, I was not allowed to go abroad … On the whole, for myself, the participation in the Atomic project proved to be a positive factor … it may have even saved my life."

Among Ginzburg's works concerned with acoustics, one of the first was the paper "On the Dispersion of High-Frequency Sound in Liquids" (Dokl. Akad. Nauk SSSR **36**, 9 (1942)), which was presented by Academician S.I. Vavilov. This subject was further investigated by Ginzburg in his paper "On the General Relation between Absorption and Dispersion of Sound Waves" (Usp. Fiz. Nauk **56**, 446 (1955)) and in the paper published in the first issue of *Akusticheskii Zhurnal*, which was cited at the beginning of this article. An analysis of dispersion relations for refraction indices and absorption coefficients of waves was performed in his paper co-authored with M.M. Meiman (Zh. Éxp. Teor. Fiz. **46**, 243 (1964)). Ginzburg's interest in acoustics is evidenced by his reviews of the book by I.G. Mikhailov *Propagation of Ultrasonic Waves in Liquids* (Gostekhizdat, Moscow, 1949; the review appeared in Sov. Kn., No. 6, 32 (1950)) and the book by L.D. Landau and E.M. Lifshits *Mechanics of Continua* (Gostekhizdat, Moscow, 1944; the review appeared in: Usp. Fiz. Nauk **28**, 384 (1946)). Ginzburg also published some other papers concerned with acoustic problems: "On Magnetohydrodynamic Waves in Gas" (Zh. Éxp. Teor. Fiz. **21**, 788 (1951)), "On the Second Sound, Convective Mechanism of Heat Conduction, and Exciton Excitations in Superconductors" (Zh. Éxp. Teor. Fiz. **41**, 828 (1961)), and other papers. A number of Ginzburg's papers are of general significance for physics of nonlinear waves: "On the Theory of Luxembourg–Gorki Effect (Izv. Akad. Nauk SSSR, Ser. Fiz. **12**, 293 (1948)), "On the Nonlinear Interaction of Radio Waves Propagating in Plasma" (Zh. Éxp. Teor. Fiz. **35**, 1573 (1958)), and "Nonlinear Phenomena in Plasma Placed in an Alternating Electromagnetic Field" (co-authored with A.V. Gurevich; Usp. Fiz. Nauk **70**, 201 (1960)).

However, it is fundamental physics that has always been of prime interest for Ginzburg, and his contributions to this field of science are great. Reflecting upon science and himself, Ginzburg considered the origin of his success. His conclusions are unexpected: "I think that my mathematical abilities are below average. … My memory, especially for formulas, is poor." However, "I have a kind of good nose, an understanding of physics, a tenacity and a combinatorial and associative grip. Secondly, I had an aspiration for 'inventing an effect,' for achieving something." These words show Ginzburg's typical attitude based on the strictest requirements and "true Nobel" criteria toward both himself and others.

"My only talent that I agree to acknowledge is the oratorical gift," writes Ginzburg. This talent clearly manifested itself in his Nobel lecture given on December 8, 2003, in the Aula Magna hall of the Stockholm University. The statement he made in 1979 that "even in English, I usually hold the audience" was fully justified. The listeners appreciated not only the outstanding scientific results but also the brilliant personality of the speaker, whose contribution to the development of human intellect and modern civilization can scarcely be exaggerated.

The author of this paper is grateful to M.S. Aksent'eva for supplying the photographs.

O.V. Rudenko

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 243–254. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 298–310.

Experimental Modeling of the Processes of Active–Passive Thermoacoustic Tomography

V. A. Burov, P. I. Darialashvili, S. N. Evtukhov, and O. D. Rumyantseva

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

e-mail: burov@phys.msu.ru

Received July 25, 2003

Abstract—Experiments confirming the major results of the theoretical study of an active–passive mode of acoustic thermotomography are described. Experimental results are obtained with a setup intended for physical modeling of the processes of correlation reconstruction of the temperature, absorption, and phase velocity of sound in the object under investigation in the presence of an additionally introduced "illuminating" acoustic noise field. A possibility of reconstructing the local values of absorption and inhomogeneity of sound velocity from analyzing the correlation dependences based on difference and summary delays and also with the help of a controlled anisotropic acoustic irradiation is demonstrated. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Recent studies have demonstrated that the tendency to increase to the maximum extent the spatial resolution and information content of images in problems of acoustic thermotomography offers considerable promise for the combined utilization of data in the active–passive mode. The approach uses a detection of the fields in a correlation tomographic system due to both the intrinsic thermal radiation of an object under investigation and the scattering of external acoustic fields by this object, including the case of the simultaneous utilization of an additional thermal or quasi-random acoustic irradiation. This active–passive mode provides an opportunity to realize a unified approach to the problem and to statistically estimate both the temperature and acoustic characteristics of a medium, which will increase the diagnostic capabilities of the acoustic methods. It is possible within the framework of this approach to suggest a new scheme of acoustic correlation thermotomography of a medium that is inhomogeneous in temperature; absorption coefficient; ultrasonic phase velocity; and, possibly, density. It is demonstrated that, in the case of using an external acoustic field (of thermal origin or specially generated) and providing the required value of the accumulation factor, the proposed method makes it possible not only to estimate the degree of heating and the character of the blood supply of a tissue [1, 2], but also to perform a separate reconstruction of the spatial distributions of all listed acoustic and temperature parameters of an object under investigation [3].

Here, we present the results of experiments illustrating the major characteristic features of the process of active–passive thermotomography. For this purpose, we developed and manufactured an experimental setup, which allowed us to perform physical modeling of the correlation detection and processing of an acoustic field of a thin refractive-absorbing layer in the presence of controlled noise irradiation.

CAPABILITIES OF ACTIVE–PASSIVE ACOUSTIC TOMOGRAPHY

A theoretical consideration of the capabilities of this method is given in [3]. Here, we briefly describe the main relationships obtained in this study and generalize them to the case of anisotropic background radiation. The process of correlation measurements that lies at the basis of thermoacoustic devices is considered [4]. A simple model of an object in the form of a thin layer that has an inhomogeneity in absorption and the phase velocity of sound provides an opportunity to analyze the correlation properties of thermoacoustic radiation and scattering and then generalize them to the case of an arbitrary object. The scheme of experimental measurements (Fig. 1) coincides with the scheme considered in [3]. Identical flat receiving transducers and a thin layer under investigation, whose thickness is much smaller than the average sound wavelength λ in it, are located in a volume filled with a liquid weakly absorbing sound and bounded by absorbing walls. From the thermal radiation of the layer, the receivers separate the plane waves that propagate in the directions perpendicular to the receiver planes. The part of the wall that is located in the sensitivity zone of a receiver emits a wave φ_0 generated by the sources F_0 as a result of the fluctuation processes in the wall. In the process of propagation, this wave is scattered by the layer with a viscosity $b(\mathbf{r})$ and a phase velocity of sound $c(\mathbf{r})$ that is different from the velocity c_0 in the surrounding liquid. The density of the layer is assumed to be close to the density of the surrounding liquid $\rho_0 = \text{const}$. The elastic properties of the layer materials (rubber and a soft plastic film) are

1063-7710/04/5003-0243$26.00 © 2004 MAIK "Nauka/Interperiodica"

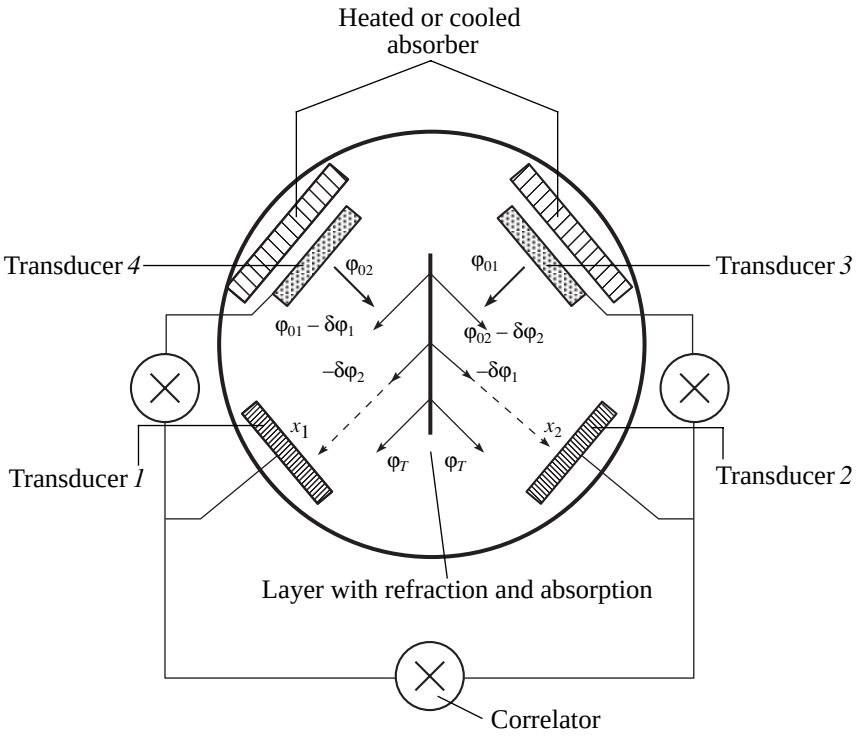


Fig. 1. Scheme of correlation measurements of the thermoacoustic fields of a thin layer.

such that the layer can be treated as a liquid one. A wave equation for the potential of the particle velocity φ in the quasi-monochromatic approximation $\sim\exp(-i\omega t)$ has the form

$$\Delta\varphi + \frac{\omega^2}{c_0^2}\varphi = \frac{ib\omega}{\rho_0 c_0^2}\Delta\varphi + \nu\varphi + F_0, \quad \text{where} \quad \nu(\mathbf{r}) = \frac{\omega^2}{c_0^2} - \frac{\omega^2}{c^2(\mathbf{r})}, \quad b = \frac{4}{3}\eta + \xi; \tag{1}$$

and η and ξ are the coefficients of shear and volume viscosity. A solution to Eq. (1) can be represented in the form $\varphi \equiv \varphi_0 - \delta\varphi$, where the scattered field $(-\delta\varphi)$ is produced by the secondary sources arising in the layer in the case of the incidence of external radiation φ_0 upon it. We examine a one-dimensional problem along the x_1 axis coinciding with the axis of directivity of receiver *1*. Its solution in the Born approximation has the form

$$\begin{aligned}\varphi(x_1) &\equiv \varphi_0(x_1) - \delta\varphi(x_1) \approx \varphi_0(x_1) \\ &- \frac{k_0^2}{2\rho_0 c_0}\int b(x')\varphi_0(x')\exp(ik_0|x_1 - x'|)dx' \\ &- \frac{i}{2k_0}\int \nu(x')\varphi_0(x')\exp(ik_0|x_1 - x'|)dx' \\ &\equiv \varphi_0(x_1) - \delta\varphi_b(x_1) - i\delta\varphi_\nu(x_1).\end{aligned} \tag{2}$$

The component $(-\delta\varphi_b)$ is antiphase to the primary field φ_0. The imaginary unit before the second integral in Eq. (2) gives a phase shift of $\pm\pi/2$ in the component $(-i\delta\varphi_\nu)$.

The correlation processing of the signals arriving from the transducers and phased by the compensation of difference or summary delays provides an opportunity to reconstruct at each spatial point $\mathbf{r}$ the following unknown characteristics of an object: the viscosity $b(\mathbf{r})$ proportional to the absorption coefficient, the inhomogeneity of the phase velocity of sound $\nu(\mathbf{r})$, the intrinsic temperature $T(\mathbf{r})$, and the temperature of background radiation $T_{bg}(\mathbf{r})$. The latter quantity $T_{bg}(\mathbf{r})$ is understood as the temperature that an absolute acoustic blackbody should have when positioned at a given point $\mathbf{r}$ and being in the state of equilibrium with the surrounding acoustic radiation. The reconstruction is based on the analysis of relationships between the thermoacoustic fields radiated directly by an element of a heated absorbing object and the external field (from other similar sources) scattered by the same element.

In the case of the mode of phasing by difference delays, the thin layer is oriented parallel to the bisectrix of the angle between the planes of receiving transducers *1* and *2* in Fig. 1. Mutually uncorrelated primary (background) waves φ_{01} and φ_{02} from the corresponding parts of the basin walls arrive at the layer lying within the region of intersection of the sensitivity zones of the receivers. After the scattering of these waves from the layer, the total field $(\varphi_{01} - \delta\varphi_1)$ and the scat-

tered field ($-\delta\varphi_1$) propagate in the directions symmetric with respect to the layer and combine with ($-\delta\varphi_2$) and ($\varphi_{02} - \delta\varphi_2$). Moreover, the absorbing layer produces its own thermoacoustic radiation φ_T, which also goes to both receivers. The cross-coherence function of signals (i.e., the correlation function in the case of a complex representation of signals) is

$$\Gamma_{12}(\tau) = \langle(\varphi_{01} - \delta\varphi_1 - \delta\varphi_2 + \varphi_T)|_t \times (\varphi_{02} - \delta\varphi_2 - \delta\varphi_1 + \varphi_T)^*|_{t+\tau}\rangle, \quad (3)$$

where the angular brackets mean averaging over realizations. In the case of the delay $\tau = \bar{\tau}_{12}$ equal to the difference of propagation times from the layer to receivers *2* and *1*, taking into account the independence of signals, we obtain

$$\Gamma_{12}(\tau = \bar{\tau}_{12}) = \langle|\delta\varphi_1|^2\rangle + \langle|\delta\varphi_2|^2\rangle - \langle\varphi_{01}\delta\varphi_1^*\rangle - \langle\varphi_{02}^*\delta\varphi_2\rangle + \langle|\varphi_T|^2\rangle. \quad (4)$$

Mutual orientation of the layer and transducers *1* and *2* provides an in-phase contribution to the quantity $\Gamma_{12}(\tau = \bar{\tau}_{12})$ from all parts of the layer. If the temperature of the basin walls is the same, i.e., the background radiation is isotropic, the imaginary part of the phased function $\Gamma_{12}(\tau = \bar{\tau}_{12})$ is mutually compensated and Eq. (4) transforms to the expression

$$\Gamma_{12}(\tau = \bar{\tau}_{12}) = 2\langle|\delta\varphi_1|^2\rangle - 2\mathrm{Re}\langle\varphi_{01}\delta\varphi_1^*\rangle + \langle|\varphi_T|^2\rangle. \quad (5)$$

Thus, in the case of phasing by difference delays, the correlated signals at receivers *1* and *2* are the scattered component of the background field and the intrinsic thermal radiation.

The intrinsic thermal field φ_{0T} generated within an absorbing object undergoes multiple scattering from the fluctuations of viscosity (b) and sound velocity (v) in the same way as the field φ_0 incident on the object from outside. Therefore, by analogy with the total field $\varphi \equiv \varphi_0 - \delta\varphi$, the total intrinsic thermal field φ_T can be written in the form $\varphi_T \equiv \varphi_{0T} - \delta\varphi_T$, where $(-\delta\varphi_T)$ is the scattered component. In this case, we have

$$\langle|\varphi_T|^2\rangle = \langle|\varphi_{0T}|^2\rangle - 2\mathrm{Re}\langle\varphi_{0T}\delta\varphi_T^*\rangle + \langle|\delta\varphi_T|^2\rangle. \quad (6)$$

Representation of the scattered fields $-\delta\varphi_1$, $-\delta\varphi_2$, and $-\delta\varphi_T$ in the form of the Born–Neumann series (their principal terms are given by Eq. (2)) provides an opportunity to express Eq. (4) using the acoustic and temperature parameters b, ν, T, and T_{bg} of the thin layer. Indeed, the term $\langle|\varphi_{0T}|^2\rangle$ from Eq. (6) provides the contribution $\sim bT$ to the value of $\mathrm{Re}\Gamma_{12}(\bar{\tau}_{12})$. The series terms linear in b and ν, which are present in Eq. (4) in the term $-\langle\varphi_{01}\delta\varphi_1^*\rangle$, produce the contribution $\sim(-bT_{bg}^{(1)})$ to the value of $\mathrm{Re}\Gamma_{12}(\bar{\tau}_{12})$ and the contribution $\sim\nu T_{bg}^{(1)}$ to the value of $\mathrm{Im}\Gamma_{12}(\bar{\tau}_{12})$, where $\langle|\varphi_{01}|^2\rangle$ is proportional to the temperature $T_{bg}^{(1)}$ of the background field φ_{01}. The analogous terms of the addend $-\langle\varphi_{02}^*\delta\varphi_2\rangle$ on account of the background field φ_{02} with the temperature $T_{bg}^{(2)}$ produce the contribution $\sim(-bT_{bg}^{(2)})$ to $\mathrm{Re}\Gamma_{12}(\bar{\tau}_{12})$ (with the same sign as the contribution of the field φ_{01}) and the contribution $\sim(-\nu T_{bg}^{(2)})$ to $\mathrm{Im}\Gamma_{12}(\bar{\tau}_{12})$ (with a sign opposite to that of the contribution of φ_{01}). Thus, the principal terms in Eq. (4) have the form

$$\mathrm{Re}\Gamma_{12}(\tau = \bar{\tau}_{12}) = 0.5A(b)[T - T_{bg}^{(1)}] + 0.5A(b)[T - T_{bg}^{(2)}] = A(b)[T - (T_{bg}^{(1)} + T_{bg}^{(2)})/2]; \quad (7)$$

$$\mathrm{Im}\Gamma_{12}(\tau = \bar{\tau}_{12}) = -\mathrm{Im}\langle\varphi_{01}\delta\varphi_1^*\rangle - \mathrm{Im}\langle\varphi_{02}^*\delta\varphi_2\rangle = C(\nu)[T_{bg}^{(1)} - T_{bg}^{(2)}], \quad (8)$$

where the coefficients $A(b)$ and $C(\nu)$ are proportional to b and ν, respectively. In the case of isotropic background radiation with the temperature $T_{bg} \equiv T_{bg}^{(1)} = T_{bg}^{(2)}$, it follows from Eqs. (7) and (8) that

$$\mathrm{Re}\Gamma_{12}(\tau = \bar{\tau}_{12}) = A(b)[T - T_{bg}]; \quad \mathrm{Im}\Gamma_{12}(\tau = \bar{\tau}_{12}) = 0. \quad (9)$$

Thus, the real part of the coherence function that is phased with the layer ($\tau = \bar{\tau}_{12}$) consists of the principal term proportional to the absorption coefficient and the temperature difference ($T - T_{bg}$). Among other terms with a higher order of smallness with respect to the components b and ν, the terms $\sim\nu^{2n}T_{bg}$ produced by scattering only from the inhomogeneities of sound velocity are always mutually compensated ($\forall n = 1, 2, 3, \ldots$) even in the case of anisotropy of background radiation. This is connected with the fact that the combinations of the terms $\langle|\delta\varphi_1|^2\rangle - \mathrm{Re}\langle\varphi_{01}\delta\varphi_1^*\rangle$ and $\langle|\delta\varphi_2|^2\rangle - \mathrm{Re}\langle\varphi_{02}^*\delta\varphi_2\rangle$ are present in $\mathrm{Re}\Gamma_{12}(\bar{\tau}_{12})$. In each of these combinations, all terms of the Born–Neumann series, which contain only ν, are compensated in all orders after phasing. Therefore, in the general case, $\mathrm{Re}\Gamma_{12}(\bar{\tau}_{12})$ consists of the sum of the terms of the type of $\sim b^{m+1}\nu^{2n}(T - T_{bg})$ ($m, n = 0, 1, 2, \ldots$) containing b.

It is important to note the following issues. First, in the case of phasing by difference delays, only the contrast ($T - T_{bg}$) is important. This is connected with the fact that, in the case of an isothermal state, the correlated condition of the intrinsic thermal radiation of the layer is exactly compensated by the anti-correlated condition of the background radiation scattered from the same layer. Second, in the case of an isotropic background field, the inhomogeneity of the sound velocity

$\nu(\mathbf{r})$ cannot be detected with difference delays, since the terms $\mathrm{Im}\langle\varphi_{01}\delta\varphi_1^*\rangle$ and $\mathrm{Im}\langle\varphi_{02}^*\delta\varphi_2\rangle$ in Eq. (8) compensate each other, and the largest term $\sim b\nu^2(T - T_{bg})$ that is present in $\mathrm{Re}\Gamma_{12}(\tau_{12}^-)$ and contains $\nu(\mathbf{r})$ has the third order of smallness with respect to inhomogeneities.

However the terms of the type of $-\mathrm{Im}\langle\varphi_0\delta\varphi^*\rangle$, which have the principal term of the series $\sim\nu T_{bg}$, can be reconstructed in the case of the mode of phasing by summary delays. Additional transducers *3* and *4*, transparent to thermoacoustic radiation and sensitive to radiation from both sides, are used to implement it. They are located on the other side of the layer, opposite to transducers *1* and *2*, respectively (Fig. 1). In this case the thin layer is oriented perpendicularly to the bisectrix of the angle between the planes of receivers *3* and *2* or receivers *4* and *1*. The cross-coherence function $\Gamma_{32}(\tau)$ for the pair (*3*, *2*) has two peaks corresponding to the delays $\tau = \pm\tau_{32}^+$. The value of τ_{32}^+ is equal to the total time of propagation of the background field φ_{01} from receiver *3* to the scattering layer and farther to receiver *2*:

$$\Gamma_{32}(\tau_{32}^+) = \langle\varphi_{01}\big|_t \times (\varphi_{02} - \delta\varphi_2 - \delta\varphi_1 + \varphi_T)^*\big|_{t+\tau_{32}^+}\rangle = -\langle\varphi_{01}\big|_t \times \delta\varphi_1^*\Big|_{t+\tau_{32}^+}\rangle. \quad (10)$$

The maximum of $\Gamma_{32}(-\tau_{32}^+)$ corresponds to the propagation of the background field in the opposite direction. In the case of an isotropic background field, we have $\Gamma_{32}(-\tau_{32}^+) = \Gamma_{32}^*(\tau_{32}^+)$.

In the case of phasing by summary delays, the correlation of signals at the receiving transducers arises only due to the external background field and corresponding scattered component. The intrinsic thermal radiation φ_T of the absorbing component of the layer is uncorrelated at the receivers. An inhomogeneity of sound velocity does not produce its own thermal radiation. Therefore, in Eq. (10), there is nothing to compensate for the wave $(-\delta\varphi_1)$ that is scattered by the layer. Its power is proportional to the power of the primary background field, i.e., the temperature $T_{bg}(\mathbf{r})$. In this case, the principal terms of the Born–Neumann series of the correlated fields $-\langle\varphi_{01}\delta\varphi_1^*\rangle$ in Eq. (10) produce two contributions for $\mathrm{Re}\Gamma_{32}(\tau_{32}^+)$, namely, the linear contribution $\sim(-bT_{bg})$ from the absorption and the quadratic contribution $\sim(-\nu^2)T_{bg}$ from the inhomogeneity of phase velocity, and one contribution for $\mathrm{Im}\Gamma_{32}(\tau_{32}^+)$, namely, the linear contribution $\sim\nu T_{bg}$. Then, we can write

$$\mathrm{Re}\Gamma_{32}(\tau = \pm\tau_{32}^+) = -0.5A(b)T_{bg} + B(\nu^2)T_{bg}; \quad \mathrm{Im}\Gamma_{32}(\tau = \pm\tau_{32}^+) = C(\nu)T_{bg}; \quad (11)$$

$T_{bg} = T_{bg}^{(1)}$ at the delay $\tau = \tau_{32}^+$. The coefficient $B(\nu^2)$ is proportional to ν^2. Unlike the case of difference delays, here the terms $\sim\nu^{2n}T_{bg}$ $(n = 1, 2, 3, \ldots)$ in $\mathrm{Re}\Gamma_{32}(\pm\tau_{32}^+)$ are not compensated, since the addends of the type of $\langle|\delta\varphi_1|^2\rangle$ are absent in Eq. (10).

Thus, the value of absorption can be reconstructed even in the conditions of isotropic background radiation in the case of signal phasing by only difference or only summary delays. At the same time, in the case of an isotropic background field, the inhomogeneity of phase velocity can only be reconstructed in the case of summary delays. However, the implementation of the scheme of correlation measurements based on summary delays is sufficiently difficult, since it requires utilization of an antenna array made of sound-transparent receiving transducers. Therefore, separate reconstruction of both scattering components $b(\mathbf{r})$ and $\nu(\mathbf{r})$ in a scheme with difference delays in the case of utilization of anisotropic background radiation seems to be promising for real tomographic systems. This is connected with the appearance of a nonzero imaginary part of a phased coherence function with the principal term proportional to $\nu(\mathbf{r})$. Indeed, if the basin walls are heated nonuniformly in such a way that the background fields φ_{01} and φ_{02} have different temperatures $T_{bg}^{(1)} \neq T_{bg}^{(2)}$, then, according to Eq. (8), $\mathrm{Im}\Gamma_{12}(\tau_{12}^-) \neq 0$. Let us consider the two simplest versions of realization of background radiation at a fixed intrinsic temperature T of a thin layer under investigation. In the first version, only the temperature of the field φ_{01} differs from T, T: $T_{bg}^{(1)} = T + \delta T^{(1)}$ and $T_{bg}^{(2)} = T$. In this case, for the coherence function $\Gamma_{12} \equiv \Gamma_{12}^{(\mathrm{I})}$ from Eqs. (7) and (8) it follows that

$$\mathrm{Re}\Gamma_{12}^{(\mathrm{I})}(\tau = \tau_{12}^-) = 0.5A(b)[T - T_{bg}^{(1)}] = -0.5A(b)\delta T^{(1)}; \quad \mathrm{Im}\Gamma_{12}^{(\mathrm{I})}(\tau = \tau_{12}^-) = C(\nu)\delta T^{(1)}. \quad (12)$$

In the second version, on the contrary, the temperature of the field φ_{02} differs from T: $T_{bg}^{(1)} = T$ and $T_{bg}^{(2)} = T + \delta T^{(2)}$. In this case, we have for $\Gamma_{12} \equiv \Gamma_{12}^{(\mathrm{II})}$:

$$\mathrm{Re}\Gamma_{12}^{(\mathrm{II})}(\tau = \tau_{12}^-) = 0.5A(b)[T - T_{bg}^{(2)}] = -0.5A(b)\delta T^{(2)}; \quad \mathrm{Im}\Gamma_{12}^{(\mathrm{II})}(\tau = \tau_{12}^-) = -C(\nu)\delta T^{(2)}. \quad (13)$$

Therefore, in the case of anisotropic background radiation, the principal term of the function $\mathrm{Re}\Gamma_{12}(\tau_{12}^-)$ is proportional to the absorption coefficient, and the principal term of the function $\mathrm{Im}\Gamma_{12}(\tau_{12}^-)$ is proportional to the inhomogeneity of sound velocity. These properties provide an opportunity to reconstruct both b and ν, while working only with difference delays. The effect can be amplified and the precision of reconstruction can be increased by using summary and difference combi-

nations of Eqs. (12) and (13). The sum $\Gamma_{12}^{(I)} + \Gamma_{12}^{(II)}$ coincides with the result of Eqs. (7) and (8) in the case of a simultaneous irradiation $T_{bg}^{(1)} = T + \delta T^{(1)}$ and $T_{bg}^{(2)} = T + \delta T^{(2)}$:

$$\mathrm{Re}\Gamma_{12}^{(I)}(\tau_{12}^{-}) + \mathrm{Re}\Gamma_{12}^{(II)}(\tau_{12}^{-}) = -A(b)[\delta T^{(1)} + \delta T^{(2)}]/2;$$

$$\mathrm{Im}\Gamma_{12}^{(I)}(\tau_{12}^{-}) + \mathrm{Im}\Gamma_{12}^{(II)}(\tau_{12}^{-}) = C(\nu)[\delta T^{(1)} - \delta T^{(2)}].$$

The difference $\Gamma_{12}^{(I)} - \Gamma_{12}^{(II)}$ yields

$$\mathrm{Re}\Gamma_{12}^{(I)}(\tau_{12}^{-}) - \mathrm{Re}\Gamma_{12}^{(II)}(\tau_{12}^{-}) = -A(b)[\delta T^{(1)} - \delta T^{(2)}]/2;$$

$$\mathrm{Im}\Gamma_{12}^{(I)}(\tau_{12}^{-}) - \mathrm{Im}\Gamma_{12}^{(II)}(\tau_{12}^{-}) = C(\nu)[\delta T^{(1)} + \delta T^{(2)}].$$

The particular case of $\delta T \equiv \delta T^{(1)} = \delta T^{(2)}$ provides an opportunity to demonstrate the advantages of the combinations $\Gamma_{12}^{(I)} \pm \Gamma_{12}^{(II)}$:

$$\begin{aligned} \mathrm{Re}\Gamma_{12}^{(I)}(\tau_{12}^{-}) + \mathrm{Re}\Gamma_{12}^{(II)}(\tau_{12}^{-}) &= -A(b)\delta T, \\ \mathrm{Im}\Gamma_{12}^{(I)}(\tau_{12}^{-}) + \mathrm{Im}\Gamma_{12}^{(II)}(\tau_{12}^{-}) &= 0; \end{aligned} \tag{14}$$

$$\begin{aligned} \mathrm{Re}\Gamma_{12}^{(I)}(\tau_{12}^{-}) - \mathrm{Re}\Gamma_{12}^{(II)}(\tau_{12}^{-}) &= 0, \\ \mathrm{Im}\Gamma_{12}^{(I)}(\tau_{12}^{-}) - \mathrm{Im}\Gamma_{12}^{(II)}(\tau_{12}^{-}) &= 2C(\nu)\delta T. \end{aligned} \tag{15}$$

The real part of the summary combination in Eq. (14) separates the contribution from the absorbing component of the layer by doubling it. In this case, the contribution of the inhomogeneity of sound velocity to the imaginary part is compensated. Conversely, the imaginary part of the difference combination in Eq. (15) separates and doubles the contribution from the inhomogeneity of sound velocity. At the same time, the contribution of the absorbing component to the real part is compensated.

The correlation properties of the fields considered using the example of a thin layer are generalized to the case of an arbitrary refractive-absorbing object. In the case of a correlation system of a circular receiving grid, the signals of each pair (i, k) of transducers form a corresponding element of the coherence matrix $\Gamma_{ik}(\tau)$. These elements are phased to the fixed point $\mathbf{r}_j$ of the object by compensation of the difference ($\tau = \tau_{ik;\,j}^{-}$) or sum ($\tau = \tau_{ik;\,j}^{+}$) of the times of propagation from this point to receivers k and i. The values of $\Gamma_{ik}(\tau_{ik;\,j}^{-})$ or $\Gamma_{ik}(\tau_{ik;\,j}^{+})$ are further summed over all pairs (i, k) with certain weights, which finally leads to the phased values of the coherence function $\Gamma_{\Sigma}^{-}(\mathbf{r}_j)$ or $\Gamma_{\Sigma}^{+}(\mathbf{r}_j)$, respectively. The cross-influence of Γ_{Σ}^{-} and Γ_{Σ}^{+} in the vicinity of the phasing point is almost absent. The properties of the functions $\Gamma_{\Sigma}^{-}(\mathbf{r})$ and $\Gamma_{\Sigma}^{+}(\mathbf{r})$ are analogous to the properties of the functions $\Gamma_{12}(\tau_{12}^{-})$ and $\Gamma_{32}(\pm\tau_{32}^{+})$ in the case of a thin layer.

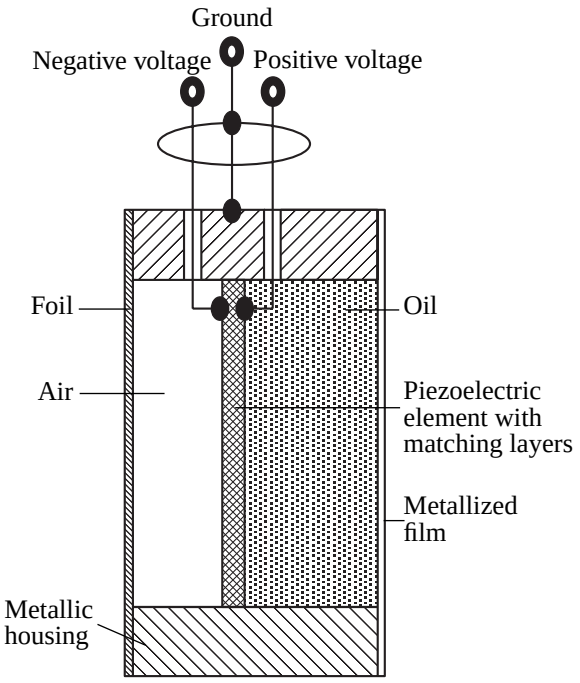


Fig. 2. Scheme of a transducer.

The values of $\mathrm{Re}\Gamma_{\Sigma}^{-}(\mathbf{r})$ and $\mathrm{Im}\Gamma_{\Sigma}^{+}(\mathbf{r})$ make it possible to reconstruct the unknown local values of $b(\mathbf{r})$, $\nu(\mathbf{r})$, $T(\mathbf{r})$, and $T_{bg}(\mathbf{r})$ for each resolution element of a region under investigation. In this case, the real part $\mathrm{Re}\Gamma_{\Sigma}^{+}(\mathbf{r})$ does not contain new information. To reconstruct separately all indicated parameters, it is necessary to conduct measurements at two different background temperatures $T_{bg}(\mathbf{r})$. One distribution of background radiation corresponds to the natural state of a system under investigation (a biological tissue), i.e., measurements in a purely passive mode. The other distribution is provided by the introduction of additional active sources into a tomographic system, which are excited by random or quasi-random signal generators and produce a field (irradiation) similar to a thermoacoustic field. This is a kind of an active mode of thermotomography. In this case, the aforementioned relationship of $\mathrm{Re}\Gamma_{\Sigma}^{-}(\mathbf{r}) \sim b(\mathbf{r})[T(\mathbf{r}) - T_{bg}(\mathbf{r})]$ and $\mathrm{Im}\Gamma_{\Sigma}^{+}(\mathbf{r}) \sim \nu(\mathbf{r})T_{bg}(\mathbf{r})$ at known values of $\mathrm{Re}\Gamma_{\Sigma}^{-}(\mathbf{r})$ and $\mathrm{Im}\Gamma_{\Sigma}^{+}(\mathbf{r})$ leads to a system of equations with respect to $b(\mathbf{r})$, $\nu(\mathbf{r})$, $T(\mathbf{r})$, and both unknown distributions $T_{bg}(\mathbf{r})$. Taking into account only the principal terms in expressions for $\Gamma_{\Sigma}^{\pm}(\mathbf{r})$ is quite sufficient because of the small size of the resolution element. Estimation of the purely refractive term $\gamma(\mathbf{r}) \sim \nu^2(\mathbf{r})$ considered in [3] is unnecessary, since the term $\gamma(\mathbf{r})$ is compensated in $\mathrm{Re}\Gamma_{\Sigma}^{-}(\mathbf{r})$. This simplifies the procedure of reconstruction of parameters $b(\mathbf{r})$ and $\nu(\mathbf{r})$ and increases its accuracy.

The possibility of using the artificial anisotropy of background radiation can be generalized to the case of an arbitrary object and a complex antenna system. In the case of a circular array, each of the terms $\Gamma_{ik}(\bar{\tau}_{ik;j})$ consists of the sum or difference of two terms $\Gamma_{ik}^{(\mathrm{I})}(\bar{\tau}_{ik;j})$ and $\Gamma_{ik}^{(\mathrm{II})}(\bar{\tau}_{ik;j})$, which correspond to the correlation of signals at the same fixed pair of receivers (i, k) but for two different versions of background irradiation in the directions of these receivers. The irradiation (i.e., radiation corresponding to the temperature different from the equilibrium temperature of the system) is produced sequentially by two sources of background radiation, which are located opposite the ith and kth receivers on the extensions of the rays issuing from the respective receiver and passing through the reconstructed element of the object with its center at the point $\mathbf{r}_j$. The situation is analogous to that considered in Fig. 1 with the thin layer being replaced by a resolution element and with the flat receivers being replaced by point receivers. The first term $\Gamma_{ik}^{(\mathrm{I})}\bar{\tau}_{ik;j}$ is formed by the direct irradiation of the ith receiver and the simultaneous detection of the scattered field (correlated with the signal at the ith receiver) by the kth receiver. In this measurement, nonequilibrium irradiation of the kth receiver from the opposite source is absent, and the direct field of the source opposite to the ith receiver is decorrelated at the kth receiver due to a time shift. The second term $\Gamma_{ik}^{(\mathrm{II})}(\bar{\tau}_{ik;j})$ is formed by the analogous direct irradiation of the kth receiver and the reception of the scattered field by the ith receiver. The result of combining the values of $\Gamma_{ik}^{(\mathrm{I})}(\bar{\tau}_{ik;j})$ and $\Gamma_{ik}^{(\mathrm{II})}(\bar{\tau}_{ik;j})$ contributes to $\mathrm{Re}\,\bar{\Gamma}_{\Sigma}(\mathbf{r}_j)$ providing an opportunity to reconstruct the absorption at a given point by taking into account all pairs (i, k). The difference combination of $\Gamma_{ik}^{(\mathrm{I})}(\bar{\tau}_{ik;j})$ and $\Gamma_{ik}^{(\mathrm{II})}(\bar{\tau}_{ik;j})$ contributes to $\mathrm{Im}\,\bar{\Gamma}_{\Sigma}(\mathbf{r}_j)$ and reconstructs the velocity inhomogeneity. Since two versions of background irradiation are determined by the point $\mathbf{r}_j$ under reconstruction and the pair (i, k), it is possible to provide them for all combinations of the parameters ($\mathbf{r}_j$, i, k) only using the nonstationary mode of active radiation, when, for example, a background source rotates in the course of the correlation measurements.

MODELING SETUP

The setup for physical modeling of thermoacoustic processes that was used for correlation processing of signals in the modes of both difference and summary delays is schematically represented in Fig. 1. Flat receiving transducers (Fig. 2) with sufficiently high electroacoustic parameters and good noise immunity were used. Therefore, the major contribution to the input noise signal was made by the acoustic noise of the medium and not by the fluctuation processes associated with losses in transducers. The sensitive element is a piezoceramic plate shaped as a disk with a diameter of 30 mm, which has a double-layer matching with the medium. The plate was positioned in a cylindrical brass housing dividing its volume into two parts. The cylinder end on one side was covered by copper foil and the space between the foil and the piezoceramics remained filled with air. This provided one-side sensitivity of the transducer, which increased the signal-to-noise ratio at the amplifier output. A polymer film with one-side metallization, which transmitted acoustic signals, was fixed on the sensitive side of the cylinder. This film with a thickness of 30 μm had an acoustic impedance value close to that of water. The space between the film and the piezoceramics was filled with silicon oil. The electric leads from the piezoceramics were connected to a differential transformer input. This design and connection considerably reduced the pickup noise of amplifiers. Special attention was given to the quality and identity of transducers in the course of their manufacture.

A pair of transducers with an operating frequency band of 890 ± 100 kHz, static capacity of ~1150 pF, impedance at the resonance frequency of 650 Ω, and sensitivity of 50 μV/Pa was used in the experiment. The input–output transfer factor for the pair of transducers at the resonance frequency is equal to 0.96 and at the boundary frequencies, to 0.59. The active part of the impedance of the receiving transducers is 400 ± 30 Ω within the frequency range 680–850 kHz. The impedance grows outside this range and is equal to 500 Ω at a frequency of 900 kHz and 1500 Ω at 1 MHz, i.e., at the upper boundary of the frequency range used.

Amplifiers were designed according to a totally symmetric paraphase circuit and were well protected against electromagnetic pickup. A high value of signal–to–internal noise ratio and small external pickup were provided by the developed and precisely manufactured scheme of amplifiers using discrete elements and transistors with a low noise factor value. The input and output of amplifiers had symmetric transformer decouplings. Utilization of two parallel amplification channels operating in antiphase in each amplifier, direct-current elements of power supply separate for each amplifier, and precise screening of circuits provided an opportunity to reduce by many times the parasitic electric pickup. The resonance frequency of the amplifiers was equal to 890 kHz with a total frequency band of about 250 kHz. The gain factor was ~350. The measured value of the noise factor was 1.5–2 dB. The effective noise resistance at the short-circuited amplifier input, which was measured within the operating frequency band and reduced to the input of the first amplifier stage (i.e., to transistor bases and not to the transformer input of an amplifier), was equal to ~8 Ω. The phase identity of the amplifiers was monitored within the whole operating frequency range. (We observed deviations within the range of 3°–5°.) The parameters of the developed transducers and amplifiers provided

an opportunity to conduct correlation measurements with a storage time of up to 60 s at a low level of correlation peak from external noise at $\tau = 0$. In the case of a storage time greater than 60 s, the level of this peak increases, although the setup permits any preset storage time up to tens of minutes.

The setup was positioned in a basin filled with water. The basin was a metallic cylindrical tank with a diameter of 325 mm and a height of 120 mm. The internal walls and bottom of the basin were covered with acoustically absorbing rubber. The thickness of the rubber layer was 17 mm on the side walls and 5 mm at the bottom. As a result, the amplitude coefficient of reflection from the walls was equal to ~0.2. Positions of all transducers and a thin layer under investigation were set and controlled with the help of independently moving rotation devices, which were fixed to the massive lid of the basin.

Digitized signals from the amplifiers were further subjected to correlation processing. It was necessary to conduct measurements for tens of seconds to obtain a sufficient statistical data level providing the necessary accumulation factor [3]. For this purpose, a scheme of digitizing with a frequency of 2.5 MHz and intermediate operative processing of signals coming in portions from two receivers was implemented. Each frame consisted of two digitized signals containing 1024 time samples each. For processing we used a standard EZ-KIT Lite card on the basis of an ADSP 21061 (Sharc) signal processor by the Analog Devices company with a clock rate of 40 MHz and also a specially manufactured card with a twelve-digit A/D converter and Altera programmable crystal (by Analog Devices), which played the role of the controller of data transfer from the A/D converter to the internal memory of the Sharc processor. With this scheme, it was possible to realize a quasi-continuous mode of signal reception: 2.5 seconds of processor operation, during which parallel correlation processing of digitized signals in a frame was performed, corresponded to one second of data acquisition. The time spectra of each of two signals in a current frame were calculated. After that, the spectra were multiplied. As the next frames were processed, a summation (averaging over frames) of spectrum products was conducted. Only the averaged product was stored, which allowed us to save memory. Processing was concluded by performing the inverse Fourier transformation of the average spectrum product, which determined the cross-correlation function. It is necessary to note that correlation processing of the real signals measured in the experiment was performed. A transition to complex representation of signals was not performed, since at this stage it was unnecessary for illustrating the basic properties of the elements of the correlation matrix. Therefore, in the figures in this paper, we present (in arbitrary units) the correlation functions of real signals. Real or imaginary parts of the coherence function correspond to these correlation functions (depending on the phase shift of their carrier frequency). A complex representation is more convenient algorithmically in the problems of thermotomography for a separate quantitative evaluation of absorption and phase velocity of sound in an arbitrary object from the real or imaginary part of the phased coherence function.

A separate problem in the course of studies was the selection of the material for a model thin layer. We used a sheet of rubber with a thickness of 0.45 mm (a medical rubber bandage) as a thin layer with strong absorption and, at the same time, a small contrast of phase velocity of sound. The phase velocity of sound in it only exceeded the sound velocity in water at ~24°C by ~30 m/s. Absorption in the operating frequency band was about 19 dB/cm, which led to absorption in the layer equal to ~0.8 dB. In what follows, we consider this layer as an absorbing one, since the absorbing component in it is much stronger than the refractive component. Indeed, at a frequency of about 1 MHz, the contributions of inhomogeneity of absorption (~2 dB/cm) and inhomogeneity of sound velocity (~10 m/s) to the scattered field are approximately equal. The second sample was considered as a refractive layer with a large contrast of phase velocity. It was an almost nonabsorbing polymer film with a thickness of ~0.15 mm. The sound velocity in it was higher than the sound velocity in water by ~350 m/s. Both samples had a surface area of 60×60 mm. Utilization of layers, where one of the components determining the scattering is most pronounced, was intended to illustrate experimentally the individual contribution of each component to the coherence function.

The layer was positioned in a small volume with the dimensions $70 \times 70 \times 20$ mm separated from the remaining volume of the basin by thin films to heat the layer up to the preset temperature T and further maintain it for several tens of seconds. This heated volume was then positioned (according to the "matryoshka" principle) into an intermediate volume with the dimensions $80 \times 80 \times 30$ mm. This structure was rigidly fixed to a common base providing an opportunity to rotate it together with the layer as a whole. In the course of the experiment, the volume where the layer was positioned, was filled with hot water, and water with a temperature slightly lower or equal to the temperature of water in the basin was poured into the second volume. This technique improves the thermal insulation of water surrounding the layer under study, since convectional heat transfer is eliminated.

EXPERIMENTAL TECHNIQUE AND RESULTS

At the stage of the preliminary adjustment of the setup, we performed precise angular orientation of the layer and receiving transducers. For adjustment in the case of measurements with difference delays, we used an additional flat piezoceramic transducer positioned into the sensitivity zone of one of receivers. In the case

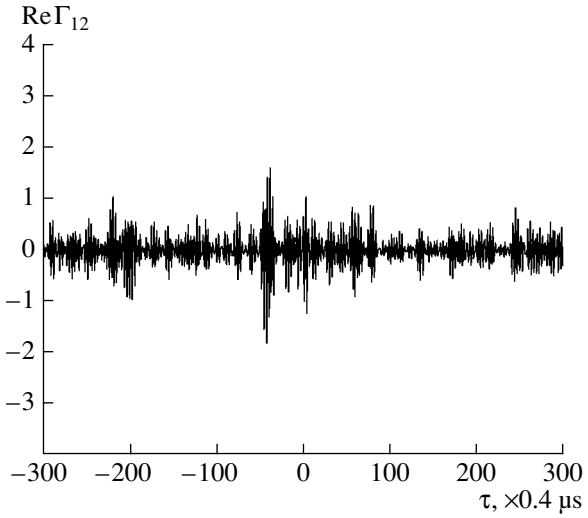


Fig. 3. Cross-coherence function of intrinsic radiation of a thin heated absorbing layer. The peak corresponds to a difference delay.

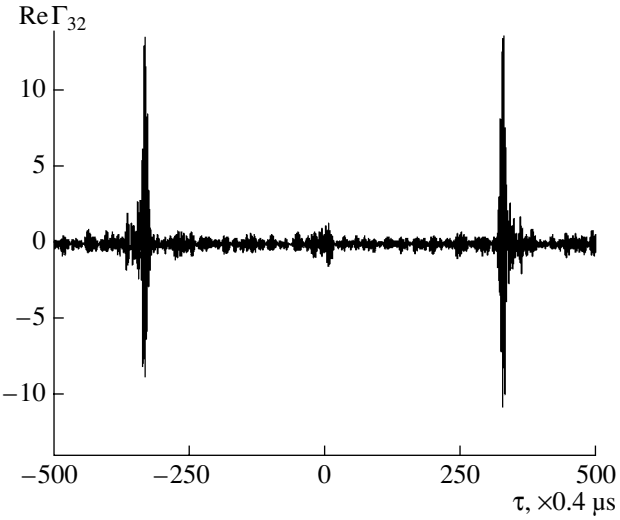


Fig. 4. Cross-coherence function of the background field reflected from a thin absorbing layer in the isothermal case. The two peaks correspond to summary delays.

of summary delays, an additional transducer was unnecessary, since here one of the transducers also played the role of a radiator.

Initially, condenser type sound-transparent transducers, which were excited by noise generators, and later (including the experiments discussed further) just pieces of thick rubber with the dimensions $70 \times 110 \times 17$ mm (similar rubber covered the internal walls of the basin) were used as the sources of irradiation, i.e., nonequilibrium background radiation. Before measurements, the rubber pieces were either heated up to a temperature of 70–80°C or cooled down to 0–3°C and positioned during measurements near the basin walls opposite receiving transducers *1* and *2*. The positions of these rubber pieces are shown in Fig. 1. In such a way, the background temperature T_{bg} was changed in comparison with the equilibrium room temperature $T_0 \approx$ 24°C of water and the other system elements. In the case of symmetric irradiation, we used two rubber pieces equally heated or cooled and positioned into the sensitivity zones of both transducers. To produce asymmetric irradiation, a piece of rubber was positioned opposite to one of the transducers, and in this case the basin wall that had the equilibrium temperature of the system got into the sensitivity zone of another transducer.

The first result presented concerns the detection of the intrinsic acoustic radiation of a thin absorbing layer. The layer was positioned into the aforementioned structure with water heated up to $T \approx (70 - 80)$°C. The background room temperature of water $T_{bg} = T_0 \approx 24$°C in the basin almost did not change. The time of correlation storage in the discussed experiments was 10 seconds (if a different time is not indicated additionally). Mutual positions of the layer and the receiving transducers corresponded to measurements by receivers *1* and *2* in Fig. 1 (the layer was oriented parallel to the bisectrix of the angle between the receiver planes). The correlation function of the intrinsic radiation of a thin layer, which corresponds to the real part of the coherence function $\mathrm{Re}\Gamma_{12}(\tau)$, is shown in Fig. 3. The maximum of the function $\mathrm{Re}\Gamma_{12}(\tau)$ is reached at the delay $\tau = \tau_{12}^{-}$ equal to the difference of times of signal propagation to receivers *2* and *1*, which agrees well with Eq. (9).

In the next experiment, the thin layer was oriented perpendicularly to the bisectrix of the angle between the receiver planes, which corresponded to measuring by the pair of receivers (*3*, *2*) or (*4*, *1*) in Fig. 1. All elements of the setup had the same room temperature. In this case, even in the isothermal mode, two peaks of the cross-coherence function $\Gamma_{32}(\tau)$ are observed, which correspond to phasing of the layer-reflected signal of the intrinsic radiation of transducers with the summary delays $\tau = \pm\tau_{32}^{+}$. According to Eq. (11), these peaks belong to the function $\mathrm{Re}\Gamma_{32}(\tau)$ in the case of a thin purely absorbing layer (Fig. 4) and the function $\mathrm{Im}\Gamma_{32}(\tau)$ in the case of a weakly absorbing layer of a refractive type (a figure for it is similar to Fig. 4). The position of one of the peaks $\tau = \tau_{32}^{+}$ corresponds to the propagation time of an acoustic noise signal (which, in this case of acoustically opaque receiving transducers, is the result of the summation of intrinsic thermal noise in a piezoelectric transducer and noise in the input circuits of an amplifier) from the third transducer to the second one. The second peak $\tau = -\tau_{32}^{+}$ corresponds to the propagation time of the signal of analogous origin from the second transducer to the third one.

To confirm illustratively Eq. (9), according to which in the case of difference delays the contribution to the

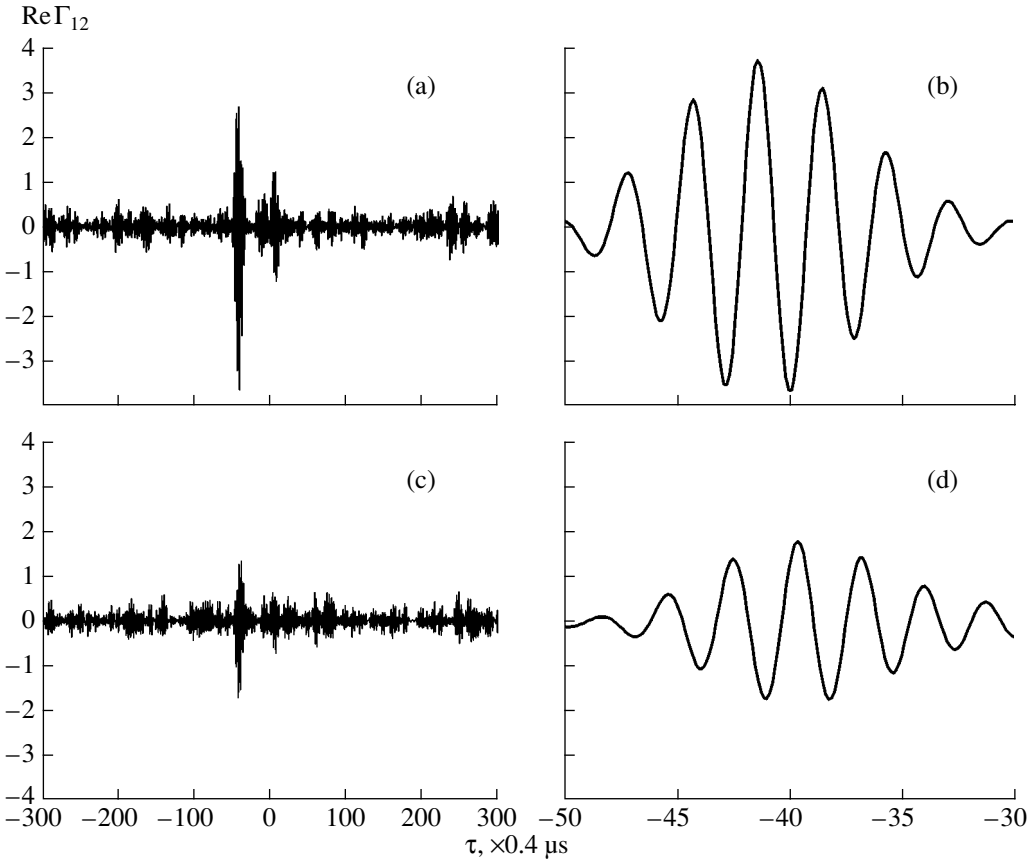


Fig. 5. Cross-coherence function for the radiation of a thin absorbing layer at a background temperature (a, b) higher and (c, d) lower than the layer temperature. The peak corresponds to a difference delay.

coherence function is produced exactly by the difference ($T - T_{bg}$) of the intrinsic temperature and the background temperature, we conducted an experiment at positive and negative temperature contrasts ($T - T_{bg}$). The scheme of the experiment corresponded to Fig. 1 with the use of the pair of transducers (*1, 2*). The background irradiation in this case was symmetric and only its temperature T_{bg} varied, which in one case was set to be higher than the intrinsic (room) temperature of the thin layer $T = T_0$ and in the second case, lower than it. Figure 5 presents the correlation functions of thermal radiation of an absorbing layer that correspond to the function $\mathrm{Re}\Gamma_{12}(\tau)$. In the cases $T_{bg} > T$ (Fig. 5a) and $T_{bg} < T$ (Fig. 5c), the background radiation was produced by heated or cooled pieces of rubber, respectively. The time of correlation storage was 20 s instead of 10 s, which was the reason for the fact that the relative noise level in Fig. 5 is $\sim\sqrt{2}$ times lower than in Figs. 3, 4, 6, and 7. The results of both experiments (Figs. 5a, 5c) agree well with Eq. (9). One can clearly see the oscillation burst of the function $\mathrm{Re}\Gamma_{12}(\tau)$ at the difference delay $\tau = \bar{\tau}_{12}$. Moreover, as it follows from Eq. (9), the sign of the function $\mathrm{Re}\Gamma_{12}(\bar{\tau}_{12})$ phased by difference delays in the two considered experiments must be opposite. This fact can be clearly observed in Figs. 5b and 5d, which show the magnified fragments of the same correlation functions after sinc-interpolation of these fragments to a thicker grid of time samples near the value of $\bar{\tau}_{12}$, which corresponds to the time sample "–39." Not only the sign but also the amplitude of $\mathrm{Re}\Gamma_{12}(\bar{\tau}_{12})$ depends on the temperature contrast. In the case of irradiation with the help of hot rubber, the temperature contrast ($T - T_{bg}$) was about –56°C, and in the case of cold rubber, about 24°C. The ratio of the absolute values of these contrasts agrees well with the ratio of the peak amplitudes of the correlation functions in Fig. 5.

The normalized values of the correlation coefficient can be obtained by division of the given values of correlation functions by the peak value of the autocorrelation function. For example, in the case of symmetric background irradiation with a temperature $T_{bg} > T$ (Fig. 5a), the peak value of the normalized correlation function is $k_{\mathrm{corr}} \cong 0.015$. On the other hand, the theoretically expected value of the correlation coefficient for an absorbing layer with a thickness l, which has the amplitude absorption coefficient α, is estimated as

$$k_{\mathrm{corr}} \approx |T - T_{bg}|[1 - \exp(-2\alpha l)]/T_{bg} \quad \text{at} \quad \alpha l \ll 1.$$

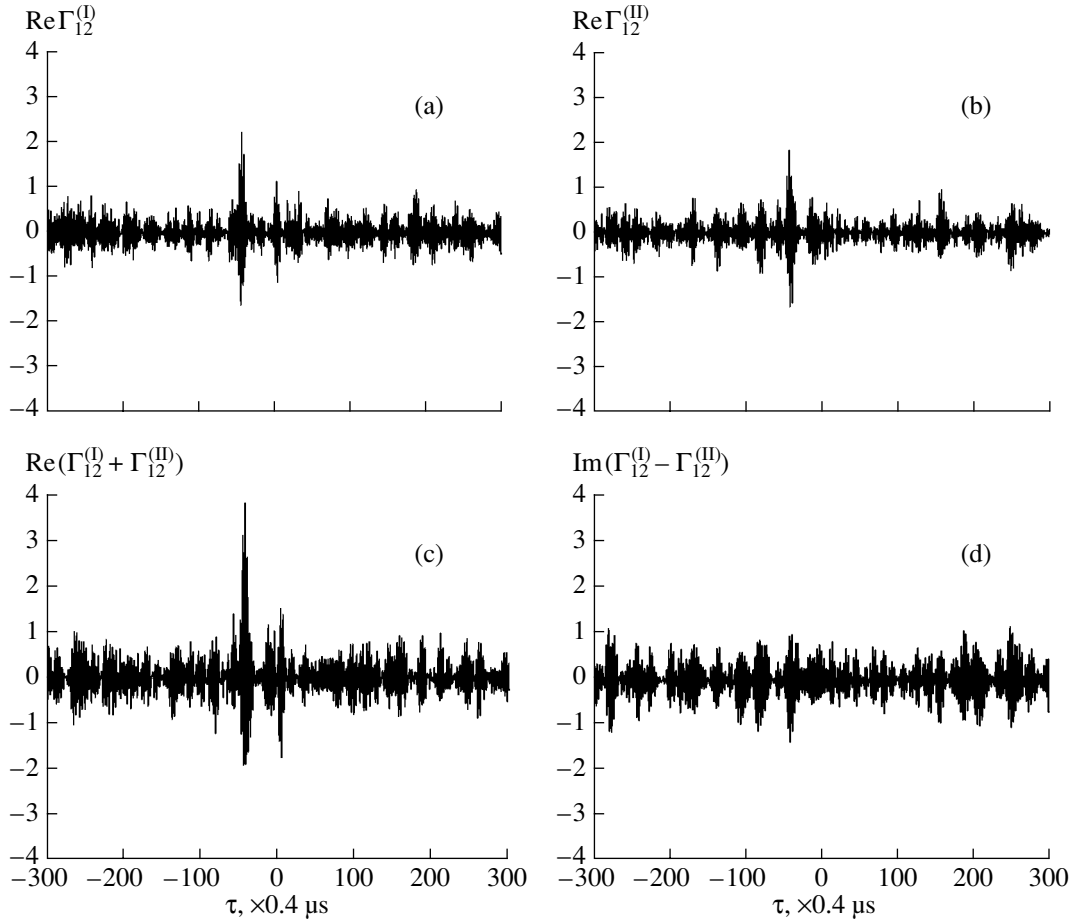


Fig. 6. Cross-coherence functions in the case of an asymmetric thermal irradiation of a thin absorbing layer in the directions of (a) only the first and (b) only the second receiver. (c) The sum and (d) difference of these functions.

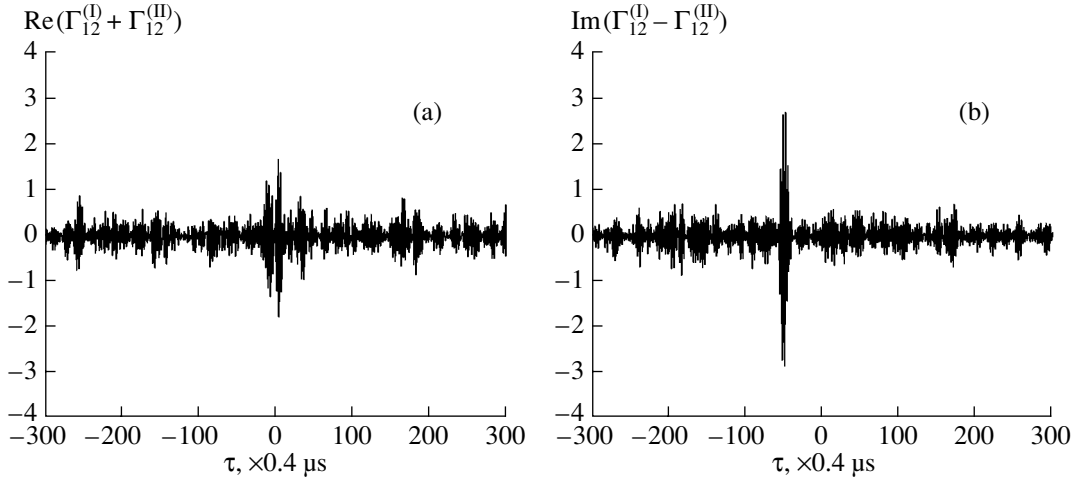


Fig. 7. (a) Sum and (b) difference of cross-coherence functions corresponding to an asymmetric irradiation of a thin refractive layer sequentially from two directions.

In the case of the experimental parameters under consideration $|T - T_{bg}| \approx 56°$, $l = 4.5 \times 10^{-4}$ m, and an absorption of 19 dB/cm, we have $k_{corr} \approx 0.03$. Taking into account the reduction of this value because of the noise factor of the receiving system and the nonideal efficiency of transducers, and also in view of the strong influence of the quality of adjustment of the experimental setup on the value of the correlation response, we can speak about a sufficiently good agreement of the experimental and theoretical values of k_{corr}.

The next experiments illustrate the capabilities of thermotomography in the case of asymmetric background radiation. The intrinsic temperature of a layer in these experiments remained equal to the room temperature, $T = T_0$. The asymmetric property of the radiation was created by thermal irradiation by one heated piece of rubber. In the first version of irradiation, the heated rubber was positioned only opposite to receiver *1*, so that the temperature of the field φ_{01} increases: $T_{bg}^{(1)} = T_0 + \delta T$ and $T_{bg}^{(2)} = T_0$. In the second version, the rubber was positioned opposite to receiver *2*: $T_{bg}^{(1)} = T_0$ and $T_{bg}^{(2)} \approx T_0 + \delta T$. Here, $\delta T^{(2)} \approx \delta T^{(1)} = \delta T$. Correlation functions for an absorbing layer in these two versions of asymmetric irradiation are given in Figs. 6a and 6b. In the case of a complex representation of signals, the functions $\mathrm{Re}\Gamma_{12}^{(\mathrm{I})}(\tau)$ and $\mathrm{Re}\Gamma_{12}^{(\mathrm{II})}(\tau)$ in Eqs. (12) and (13), respectively, have an analogous form. In this case, the functions $\mathrm{Im}\Gamma_{12}^{(\mathrm{I})}(\tau)$ and $\mathrm{Im}\Gamma_{12}^{(\mathrm{II})}(\tau)$ are shifted with respect to them to a quarter-period of the carrier frequency. The sum of correlation functions (Fig. 6c) almost reproduces the correlation function of Eq. (9) obtained in the case of symmetric irradiation of the layer. This sum corresponds to the combination $\mathrm{Re}\Gamma_{12}^{(\mathrm{I})}(\tau) + \mathrm{Re}\Gamma_{12}^{(\mathrm{II})}(\tau)$. The combination $\mathrm{Im}\Gamma_{12}^{(\mathrm{I})}(\tau) - \mathrm{Im}\Gamma_{12}^{(\mathrm{II})}(\tau)$ corresponds to the difference of correlation functions (Fig. 6d). According to Eq. (14), the summary combination phased to the layer separates the contribution of the absorbing component. Indeed, for the layer under consideration, this sum has a clearly expressed maximum at $\tau = \bar{\tau}_{12}$ (Fig. 6c). The maximum is doubled in comparison with the versions of one-side irradiation (Figs. 6a and 6b). At the same time, a small amplitude for the difference combination of Eq. (15), which carries information on the refractive component, is evidence of the almost complete absence of the velocity contrast in the layer (Fig. 6d).

The correlation functions of the refractive layer at two versions of asymmetrical irradiation have similar shapes in Figs. 6a and 6b. In this case, the correlation functions are antiphase near the value of the phasing delay $\tau = \bar{\tau}_{12}$. This fact agrees well with Eqs. (12) and (13): $\mathrm{Im}\Gamma_{12}^{(\mathrm{II})}(\bar{\tau}_{12}) = -\mathrm{Im}\Gamma_{12}^{(\mathrm{I})}(\bar{\tau}_{12})$ at $\delta T^{(1)} = \delta T^{(2)}$, and the absorbing component is absent in the layer. The latter fact is confirmed by the form of the summary correlation combination, which has no characteristic rises (Fig. 7a). The difference combination (Fig. 7b) contains a maximum at $\tau = \bar{\tau}_{12}$. It is doubled in comparison with one-side irradiation. In analogous experiments with samples that have both *b*- and ν-components, the characteristic peaks are present in both summary and difference combinations. Their values are proportional to the values of *b* and ν, respectively. Thus, in the case of signal phasing by difference delays, an opportunity to reconstruct not only the values of absorption (such opportunity exists also in the case of an isotropic background radiation) but also the inhomogeneity of the phase velocity of sound is demonstrated. This becomes practically possible only in the case of anisotropic background radiation.

It is necessary to specify some technical details of conducted modeling experiments. Their purpose at this stage was an illustration of the main theoretical results but not a quantitative evaluation of the temperature and acoustic parameters of an object and the surrounding medium. Therefore, the setup was designed according to a simplified scheme, which made it possible to compare the results quantitatively only within a single series of experiments conducted at a fixed setup adjustment. It is possible to improve the setup by increasing its precision characteristics. Such an improvement that is not accompanied by fundamental technological difficulties would provide an opportunity to quantitatively reconstruct the values of all parameters (b, ν, T, T_{bg}) of an object under investigation. In the general case of an object of arbitrary shape, it is necessary to implement a full scheme of active–passive thermoacoustic tomography, where the number of receiving transducers and the necessary accumulation time are determined by the required resolution [3].

CONCLUSIONS

Currently, the theory of the processes determining the capabilities and difficulties of the active–passive mode of thermoacoustic tomography is clear in its major aspects. Basic theoretical results have been confirmed experimentally. At this stage, understanding of the physical nature of thermoacoustic processes is sufficient for the further development of thermotomography systems, but the main difficulty in implementing such systems is the low input signal-to-noise ratio. This difficulty is of a fundamental character. It is common to all systems of noncoherent tomography, and its elimination requires antenna systems with a very large number of transducers and a long averaging time.

ACKNOWLEDGMENTS

We are grateful to K.N. Bobov for his assistance in organizing the computer support and to M.I. Rattel' for

participating in the development and manufacture of the receiving transducers.

The work was supported by a grant from the President of the Russian Federation (grant no. NSh-1575.003.2) and by the Russian Foundation for Basic Research (project no. 01-02-16282).

REFERENCES

1. A. A. Anosov, K. M. Bograchev, and V. I. Pasechnik, Akust. Zh. **44**, 725 (1998) [Acoust. Phys. **44**, 629 (1998)].
2. V. A. Burov, E. E. Kasatkina, O. D. Rumyantseva, and S. A. Filimonov, Akust. Zh. **49**, 167 (2003) [Acoust. Phys. **49**, 134 (2003)].
3. V. A. Burov, P. I. Darialashvili, and O. D. Rumyantseva, Akust. Zh. **48**, 474 (2002) [Acoust. Phys. **48**, 412 (2002)].
4. R. A. Hessemer, Jr., T. Perper, and T. Bowen, US Patent No. 4,416,552 (22 November 1983), p. 16.

Translated by M. Lyamshev

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 255–264. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 311–321.

Long-Range Sound Propagation in the Mediterranean Sea

R. A. Vadov

Andreev Acoustics Institute, Russian Academy of Sciences, ul. Shvernika 4, Moscow, 117036 Russia
e-mail: vadov@akin.ru

Received June 20, 2003

Abstract—The data of several experiments on the long-range propagation of explosion-generated and tonal sound signals are analyzed. The experiments are performed by the Acoustics Institute in the Mediterranean Sea with a fully developed sound channel. A substantial difference is observed for the propagation conditions in the western and eastern parts of the sea. This difference concerns the vertical sound speed profiles, the time structures of the sound field in the underwater sound channel, the duration of the explosion-generated signal, and the positions of the convergence zones. The experiment is compared with calculations. The observed difference in the experimental and calculated positions of the first convergence zone is explained by the imperfection of the relation used to recalculate the salinity, water temperature, and hydrostatic pressure to the sound speed. In spite of substantial difference in the propagation conditions on two 600-km paths, the experimental low-frequency attenuation coefficients on these paths (and on some shorter ones) agree well with each other for the frequency band of several kilohertz. The data are also close to those published for another 600-km path. All the paths mentioned run in different parts of the Mediterranean Sea. The frequency dependence of sound attenuation (absorption) can be well described by the relation that accounts for the absorption caused by the boron present in the sea water. *© 2004 MAIK "Nauka/Interperiodica".*

The Mediterranean Sea is a land-locked one for which a specific hydrological regime is characteristic [1–4]. Its main specificity is a moderate water interchange with other regions (the Atlantic Ocean, the Black Sea, and the Red Sea). The evaporation rate substantially exceeds the income of atmospheric precipitation and the yield of rivers. The water shortage is compensated by the surface waters of the Atlantic, which enter the sea through the Strait of Gibraltar. The evaporation-caused increase in salinity leads to an increase in the water density. The denser water comes down and gradually fills the western and eastern troughs of the Mediterranean Sea to form a uniform bulk of relatively warm high-salinity water. At depths of 250–300 m, the temperature and salinity of Mediterranean water slightly varies (12.5–13.5°C to 37.5–39‰, respectively). In summer and autumn, the warming of surface water forms an underwater sound channel (USC) with its axis at a depth of 150–300 m. The Mediterranean USC is specific in that it is concentrated near the surface (with near-bottom sound speed considerably greater than the near-surface one) and has a small thickness (with a sea depth of 3000 m or more, the "purely water" signals propagate within a near-surface layer of 1200–1700 m in thickness).

Being prolate in the latitudinal direction, the Mediterranean Sea consists of several basins. However, the Sicilian Ridge situated between Sicily and Africa divides the sea into two large zones, the western and eastern troughs, with encumbered water interchange between them. The strait over the ridge is a shallow-water area with a narrow passage whose depth is less than 400 m.

The western and eastern parts of the Mediterranean Sea are quite different in sound propagation. This difference manifests itself foremost in the vertical profiles of the sound speed and in the USC shapes. Figure 1 shows profiles $c(z)$ measured in autumn, on two 600-km paths crossing the Algeria–Provence basin and the Levant Sea, respectively. In the western region, the USC has a channel axis at depths of 50–200 m. In the eastern region, the USC axis is not as pronounced and lies at depths of 150–500 m (from 200 to 400 m the sound speed varies by no more than ±0.5–0.7 m/s). The minimal sound speeds are also considerably different in the two regions. Such a difference is caused by the isolation of these regions, the weak water interchange both between them and with the Atlantic Ocean, and their different distances from the Strait of Gibraltar.

The first experiments on sound propagation in the Mediterranean Sea were performed in the early and mid-sixties of the last century. Starting in 1963, more or less systematic studies were carried out under the NATO underwater-acoustics research projects. In the early experiments [5], the path lengths were several tens of kilometers, and they were increased up to 600 km in 1982 [6]. In recent years, much attention has been paid to the acoustic tomography of the Mediterranean waters [7].

Since 1965, the Acoustics Institute studied the intensity, time, and spatial structures of the sound field produced by a point source in both the eastern and western parts of the Mediterranean Sea. The studies on long-

1063-7710/04/5003-0255$26.00 © 2004 MAIK "Nauka/Interperiodica"

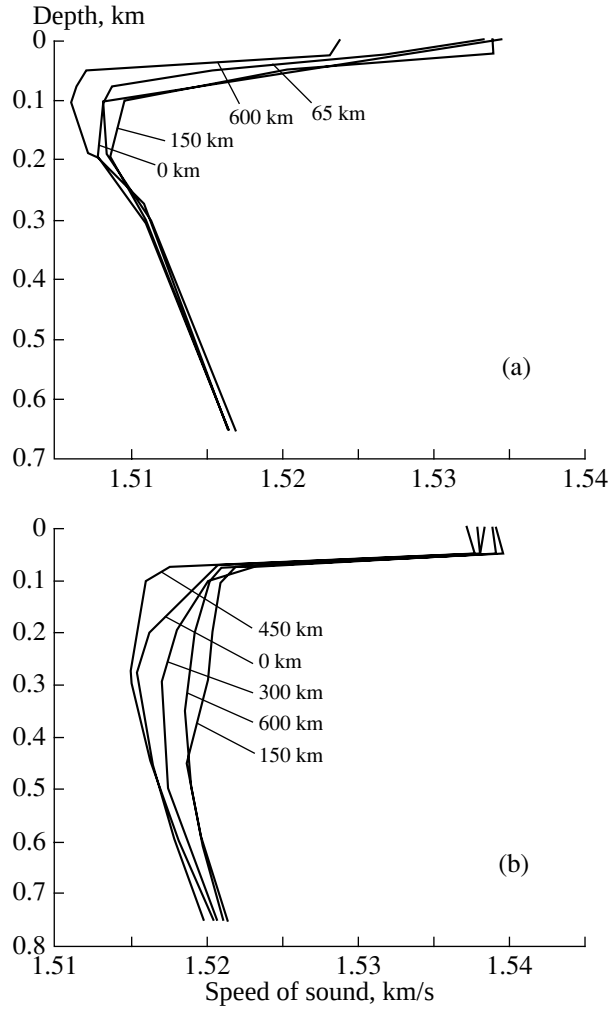


Fig. 1. Vertical sound speed profiles measured in the (a) western and (b) eastern parts of the Mediterranean Sea on the paths of long-range propagation of explosion-generated signals. The distances from the receiving vessel to the points of measuring the profile $c(z)$ are indicated to the right of each curve.

range sound propagation of explosion-generated signals (on paths longer than 600 km) were carried out in autumn (September and October), 1965, in two regions (the Algeria–Provence basin and the Levant Sea), which differed in both their water characteristics and bottom relief.

In the sixties and seventies, the high interest in the Mediterranean Sea was particularly caused by the specificity of water characteristics, namely, by higher water temperature and higher salinity in comparison with the ocean. The experimental data on sound attenuation in the Mediterranean, Black, and Baltic Seas, and the Sea of Japan, which differ considerably in water temperature and salinity, offered a good opportunity to verify the hypothesis of the boron-caused low-frequency relaxation proposed in the early seventies.

Measurements in the kiloherts frequency range were carried out using piezoceramic transducers as the sources of sound. The objective of these experiments was to study the intensity characteristics of the sound field (foremost, sound attenuation and absorption), including the vertical and horizontal structures of sound fields.

In the western part of the sea, the experiments were performed in both the spring and autumn. In spring (May), the measurements were carried out south-east of the Balearic Islands. The sea depth on the path was 2400–2800 m. The USC axis was at depths of 60–120 m. Purely water signals propagated in the water layer lying between the surface and a depth of 720 m. A towed signal source was used with a towing depth of 100–110 m. Continuous pseudonoise signals were transmitted in a 1/3-octave band around the central frequencies of 1.6 and 4 kHz. An omnidirectional receiving system was at a depth of 100 m. The longest distance between two vessels involved in the experiment was 120 km. The attenuation coefficient was estimated by measuring the deviation of the experimental sound level decay from the cylindrical law and was found to be 0.13 and 0.32 dB/km, respectively. Figure 2 illustrates the decay of the sound field level at 4 kHz, upon compensating for the cylindrical law of the geometric spread. The method of least squares was used to obtain the slope of the straight line approximating the decay at distances of 20 to 90 km. The slope of this line relative to the abscissa axis (0.32 dB/km in our case) determines the attenuation coefficient. The data of this experiment were also used to estimate the transition distance, which proved to be 2.7 km in the spring.

In autumn (the second half of November), a path south of the Balearic Islands was surveyed. With a towed sound source, semidiurnal studies were carried out to observe the sound field level in the vicinity of the first convergence zone, at distances of 25–40 km from the source. The sea depth was 2800 m on the path. The axis of the somewhat “fuzzy” USC was at a depth of 150–200 m with a discontinuity layer depth of about 60 m. The near-surface mixed water layer was about 50 m in thickness. The purely water signals propagated at depths from 50 to 1100 m.

Noise-like 1/3-octave signals around 2.5 kHz were transmitted. The experimental procedure was based on the method proposed earlier (see [8, 9]). The sound source was towed at a depth of 84–88 m, and the reception depths were 15, 200, and 600 m. Over twelve hours, the transmitting vessel crossed the first convergence zone five times.

For the reception depths of 200 and 600 m, the rms deviation of the sound field level reached 2.5–3 dB at some points of the path.

A pronounced zonal structure of the sound field only manifested itself for the reception depth of 15 m. The rms deviation of the sound field level was no higher than 2 dB within the convergence zone. The experiment-

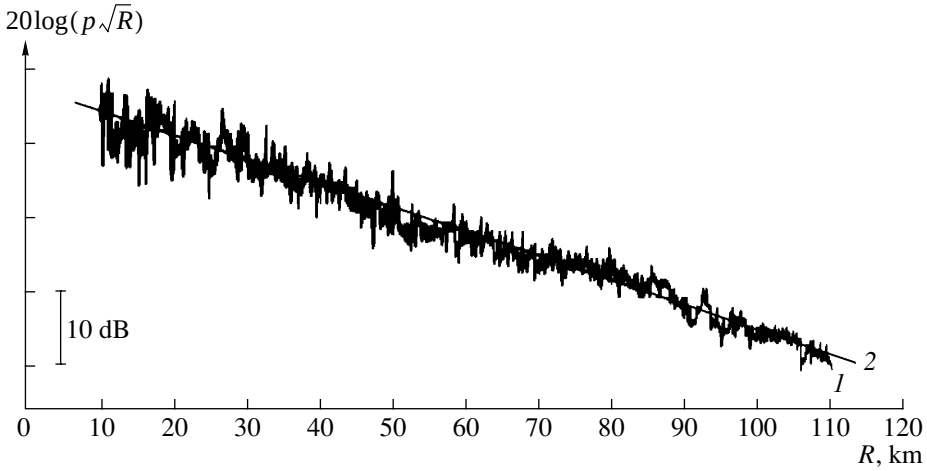


Fig. 2. Experimental sound field level decay compensated for the cylindrical spread in the underwater sound channel of the western part of the Mediterranean Sea. The sound frequency is 4 kHz.

al data were used to estimate the distance to the nearest (from the source) boundary of the convergence zone. The mean value of this distance proved to be 28.4 km, with the 195-m rms deviation of the boundary position (as estimated for five horizontal sections of the sound field). To calculate the distance from the source to the nearest boundary of the convergence zone, the computer code by Tebyakin [10] was used to yield 29.925 to 30.075 km. The computations used five profiles $c(z)$, which were measured after each of the five crossings of the convergence zone. The difference between the experimental and calculated distances to the zone was 1.5 km.

In [11], the main reason for such a difference was considered to be the ambiguity in the representation of the deep-water part of profile $c(z)$. Such an ambiguity can be caused by an inaccuracy of the formulas used to calculate the sound speed from the measured depth, temperature, and salinity. The author of this paper also performed a series of calculations. With the depth dependences of temperature and salinity measured just before the acoustic experiment, formulas proposed by different researchers were used to calculate the profiles $c(z)$. The distances to the nearest boundary of the convergence zone were determined for each of these profiles. The results obtained are summarized in Table 1.

The spread of the resulting distances to the zone proved to be about 3 km, which is two times greater than the aforementioned difference between the experiment and computations. The experimental values of the distance fall between the values corresponding to profiles $c(z)$ given by the formulas of Kuwahara, Frye, and Lovett, with a difference of 1.5–1.8 km between these estimates. Here, it is worth mentioning publication [12]. In the "Discussion" section of [12], English argues that, in the experiments of WHOI in the Mediterranean Sea, the measured distance to the convergence zone was between the estimates based on the data of Kuwahara and Wilson. From considerations by English, it is also clear that the experimental value of this distance is closer to that given by the Kuwahara formula. This result quite agrees with our calculations. It seems that none of the proposed formulas gives the correct result for the Mediterranean Sea, whose temperature and salinity differ considerably from those of the Atlantic Ocean. Thus, the formulas for calculating the sound speed from temperature and salinity must be refined.

In the eastern part of the Mediterranean Sea (the Ionian Sea, south-east of the Sicily Island), the experiments were carried out in autumn (November). There was a near-surface layer mixed down to a depth of 40–50 m, with a positive sound speed gradient close to the hydrostatic one. The seasonal thermocline, not yet destroyed by the winter vertical convection, was most pronounced at depths of 50–70 m. The mean sound speed gradient in this layer was 0.8 s^{-1}. The lower boundary of the thermocline reached the depths of 85–100 m. A fully developed USC was observed with a weakly pronounced axis at depths of 120–200 m. Purely water signals propagated within the layer occupying depths from 50 to 1350 m. The sea depth varied within 3600–3900 m with a moderately rough bottom relief. The sound source was towed at a depth of 110 m. The reception depths were 10 and 120 m.

The decay of the sound field propagating in the USC was determined from the signals received at a depth of 120 m, on the 85-km path at a frequency of 3.15 kHz and on the 100-km path at 1.6 kHz. This decay was used to estimate the attenuation coefficients, which proved to be 0.235 and 0.095 dB/km, respectively. Figure 3 shows the sound field decay compensated for the cylindrical spread law. To determine the attenuation coefficient, the part of the decay curve was used that corresponded to the distances from 20 km to the ulti-

Table 1. Distance from the source to the nearest boundary of the convergence zone: results of the sound field calculations with profiles $c(z)$ plotted for the western part of the Mediterranean Sea according to different formulas. The source and receiver depths are 86 and 15 m, respectively

S. Kuwahara (1960)	W.D. Wilson (1962)	V.A. Del Grosso (1972)	H.W. Frye (1971)	C.T. Chen (1977)	J.R. Lovett (1978)	K.V. Mackenzie (1981)	A.B. Coppens (1981)
27.775	30.475	30.325	27.525	29.675	29.325	29.775	29.975 km

mate range. In Fig. 3, a straight line is presented that approximates the decay for that path fraction.

Figure 4 shows the decay of the sound field received at a depth of 10 m with the transmitting vessel moving away from the reception point to a distance of 50 km. A continuous signal was transmitted at a frequency of 1.6 kHz. The nearest boundary of the first convergence zone was observed at a distance of 33.2–33.3 km from the source.

Two days later, the next experiment, with a broadband sound source, was performed in the same region to study the vertical structure of the sound field. The source was at a depth of 140 m. A noise-like signal was transmitted in the frequency band within 0.5–5.0 kHz. At certain distances from the source, an omnidirectional hydrophone was deployed from the receiving vessel, and the sound signal was continuously received down to a depth of 500 m. Upon filtration in 1/3-octave frequency bands, the received signals were recorded with the use of the 2307 Bruel&Kjer chart-recording instrument. The maximal distance between the vessels was 200 km.

From this experiment, the vertical structure of the sound field, the frequency dependence of this structure, and its changes in distance were estimated. In addition, data on the frequency dependence of sound attenuation were also obtained. To estimate the attenuation coefficient, the sound field levels were energetically averaged for the depths from 100 to 400 m, at each frequency and each distance. The values of the attenuation coefficient were determined by measuring the deviation of the averaged level decay from the cylindrical law. These values were 0.08, 0.2, 0.21, and 0.4 dB/km at frequencies 1.0, 2.0, 2.5, and 5.0 kHz, respectively.

Studies with explosion-generated sound signals were carried out at lower frequencies (100–1000 Hz). These studies provided the determination of both intensity and time structure of the sound field.

The long-range propagation of explosion-generated signals was studied in the western part of the Mediterranean Sea, on a path that crossed the Algeria–Provence basin from south-west to north-east and was 640 km long. The sea floor along the path was smooth (a deep-water plain), covered with a sediment layer of 1.0–1.5 km in thickness, with a low propagation velocity of

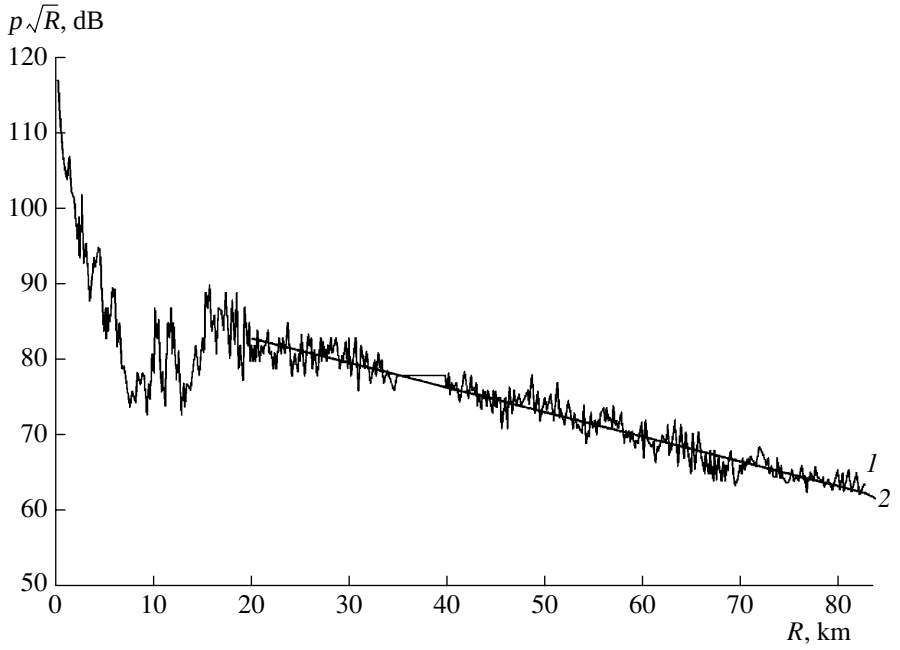


Fig. 3. Experimental sound field level decay compensated for the cylindrical spread in the underwater sound channel of the eastern part of the Mediterranean Sea. The sound frequency is 3.15 kHz.

longitudinal waves (1.7–2.5 km/s). The rocky bulk underlying the sediment layer had a velocity of 3.0–6.0 km/s for longitudinal waves [1]. The samples of the upper layer of bottom sediments contained interleaving layers of sand, siltstone, and clay. The sea depth varied from 2700 to 3000 m along the path. The experiment was carried out in autumn (end of September), at low wind speed and sea state (less than Beoufort 2). The water characteristics were rather uniform along the path. The upper mixed layer was 25 m in thickness, and the discontinuity layer covered the depths from 25 to ~75 m. The lower boundary of the discontinuity layer was at depths of 50 and 100 m in the northern and southern parts of the region, respectively. The sound speed at the slightly pronounced USC axis (100–150 m) was 1506–1509 m/s. From 200–300 m to the bottom, both the temperature and salinity of the water remained nearly constant. The sound speed variation was caused by nothing but the hydrostatic gradient, with a monotonic increase in depth. The near-surface sound speed was higher by 19–26 m/s than that at the USC axis. The sound speed near the bottom exceeded that at the USC axis by 42–45 m/s. Figure 1a shows several sound speed profiles measured along the propagation path during experimentation. A weakly perturbed sound speed field and a rather flat bottom are characteristic of the path at hand.

In this experiment, the receiving vessel drifted at 38°15′ N, 5°30′ E. The transmitting vessel went from the point about 80 km south of the Monaco Port towards the receiving one. From the transmitting vessel, 25-kg explosive charges with pressure-sensitive detonators were dropped and exploded at a depth of 300 m. The total number of charges was 134. The explosion-generated signals were received by an omnidirectional receiving system at a depth of 200 m. At the moment of dropping each charge, the distance between the vessels was determined from the signal propagation time. The received signals were magnetic-tape recorded.

Under the conditions of single-path propagation, the signal transmitted from a distance of 10–20 km or more and received in the frequency band from 10–20 Hz to 1–2 kHz consists of two short (less than 1 ms) pulses that are equal in their amplitudes and have the same signs. These pulses are produced by the shock wave and the first gas-bubble oscillation. The time interval between the pulses corresponds to the period of the first bubble oscillation, about 55 ms in our case. In the case of multipath sound propagation, a pair of such pulses corresponds to each ray in the time structure of the explosion-generated signal.

Figure 5 illustrates the range-dependent duration and time structure of a multipath signal. Here, the signals are presented that were received and recorded at distances of 49 to 644 km from the source with approximately 50-km steps in distance. All signals are normalized to their maximal magnitudes. Apart from bottom-reflected elementary signals, the total duration of a multipath signal was, in the mean, proportional to the distance with a proportionality factor of about 0.006 s/km. A multipath explosion-generated signal breaks down into three groups of signals. The first, most intense main group is formed by elementary (single-path) signals that can hardly be separated in time. The second group consists of "classical" quartets of purely water and surface-reflected signals that arrive at the receiver with some advance relative to the main group. The third group contains quartets of bottom- and surface-reflected signals that are received with some delay relative to the main group. The data presented in Fig. 5 cannot fully characterize the variation of the signal time structure in distance: the 50-km distance step is too rough. Figure 6 presents the explosion-generated signals received at distances of 50–135 km from the source with 5-km steps in distance. Here, the features of range-dependent changes in the signal structure are evident (the 5-km step proved to be sufficiently small). As the distance increases, the classical quartets of time-separated elementary (single-path) signals become distinguished from the main group of the near-axis signals. The quartet becomes more and more ahead of the group of the near-axis signals, and, finally, the quartet of purely water signals transforms into the quartet of surface-reflected ones. At distances that are multiples of about 40 km, such a quartet vanishes when the rays touch the bottom. Then, the quartet is replaced by the next one that differs from the previous quartet by one lower half-cycle of the corresponding ray.

To the right of the main signal group, a quartet of signals that are time-delayed because of reflections from the bottom (and surface) can be observed. The delay of these signals relative to the main group monotonically decreases as the distance increases. The quartet of single-bottom-reflected signals is followed by the quartet of signals with two, three, and so on bottom

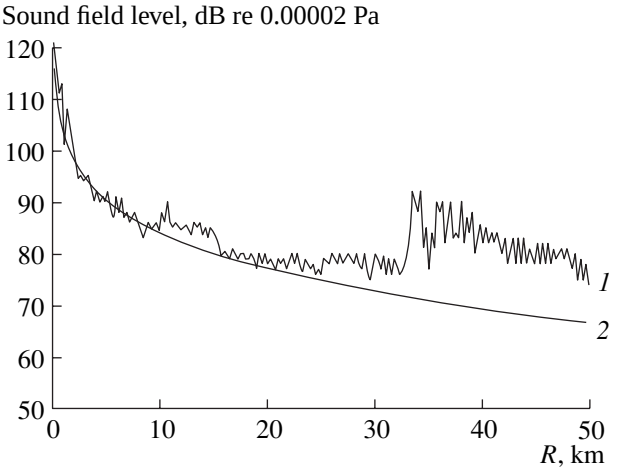


Fig. 4. Experimental sound field level decay in the eastern part of the Mediterranean Sea. The frequency is 1.6 kHz, the source towing depth is 110 m, and the reception depth is 10 m.

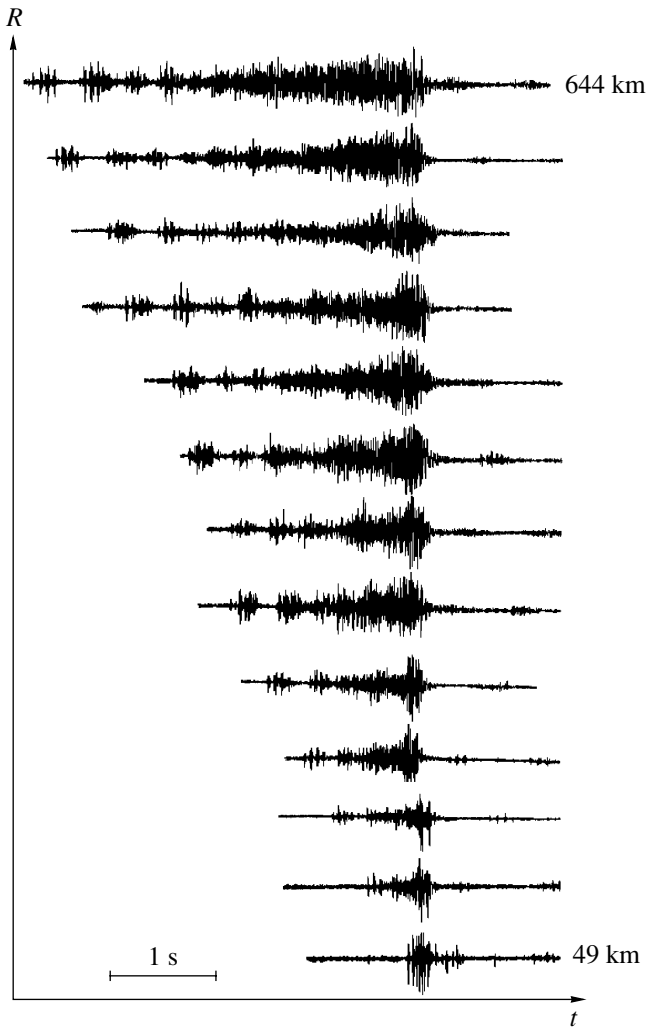


Fig. 5. Changes in the time structure of the sound field propagating in the USC of the western part of the Mediterranean Sea. The distance from the source varies from ~50 to ~650 km with steps of ~50 km. Each signal is normalized to its maximal amplitude.

reflections. In the experiment at hand, signals with multiple bottom reflections were observed at all distances, up to the ultimate one (644 km).

In [13], the reduced $t/N–R/N$ curve was proposed as a compact characteristic of the time structure for a signal propagating in the USC without touching the bottom. Here, t is the advance time relative to the signal propagating along the USC axis, N is the number of lower half-cycles (the ordinal number of the quartet), and R is the distance. Figure 7 shows the $t/N–R/N$ curve plotted for three branches of the quartets (for $N = 2$, 3, and 4 and $R = 50–135$ km) on the basis of the experimental data. Such a curve fully determines the position of time-separated quartets in the time structure of a multipath signal at an arbitrary distance from the source. The experimental data and results of calculations (for $N = 1$) concerning the analogous dependence can be found in [5]. In spite of the fact that this experiment was also performed in the western part of the Mediterranean Sea, the values of $t(R)$ reported in [5] are somewhat lower (by a factor of 1.05–1.1) than those presented in Fig. 7a. The proposed dependence $t(R)$ has the form $t_{ms} = (0.015R_{\text{km}})^3$. Our curve also follows a power function, although with a somewhat different exponent (2.9 instead of 3) and proportionality factor (0.0155 instead of 0.015).

In the eastern part of the Mediterranean Sea, the propagation path of explosion-generated signals went in a latitudinal direction (from west to east), along the northern boundary of the Levant Sea. The path was 610 km in length. The sea floor was rather rough along the path with depth drops of several hundreds of meters. The experiment was carried out in autumn (beginning of October), with weak winds and sea states of Beoufort 0–1.

The changes in water characteristics occurred mainly in the 500–600–m water layer. The near-surface mixed waters reached a depth of 25–50 m, and the temperature discontinuity layer with a sound velocity gradient up to -1 s^{-1} occupied the depths from 25–50 to 50–75 m. A rather weakly pronounced minimum of sound speed was observed at depths of 100 to 500–600 m (the changes in the sound speed were no greater than ±0.5–0.7 m/s within the layer from 200 to 400 m in depth). Deeper, the sound speed monotonically increased in depth because of the hydrostatic gradient (the water temperature and salinity remained nearly constant). Figure 1b shows the sound speed profiles measured during experimentation. The depth of the sound speed minimum varies from 250–300 to 450 m along the path. The value of the minimum is within 1515–1518 m/s.

In this experiment, the reception point was 50–60 km south-east of Cyprus Island. The transmitting vessel went a heading of about 275° away from the receiving vessel by throwing explosive charges overboard. The 25-kg charges equipped with pressure-sensitive detonators exploded at a depth of 200 m. In total, 73 charges were dropped along the path.

In this Mediterranean region, the total duration of the received multipath explosion-generated signal changed (in the mean) proportionally to the distance with a proportionality factor of about 0.0027 s/km. The multipath signal consisted of the main group of elementary (single-path) signals, which cannot be time-resolved, and the group of classical quartets consisting of purely water and surface-reflected signals, this latter group arriving with some advance relative to the main group. The quartets of bottom reflections were not observed in this experiment: bottom reflections were accompanied by strong attenuation of sound signals, particularly because of the high roughness of the sea floor.

Figure 8 clearly shows the character of variation of the time structure of signals in distance. This figure pre-

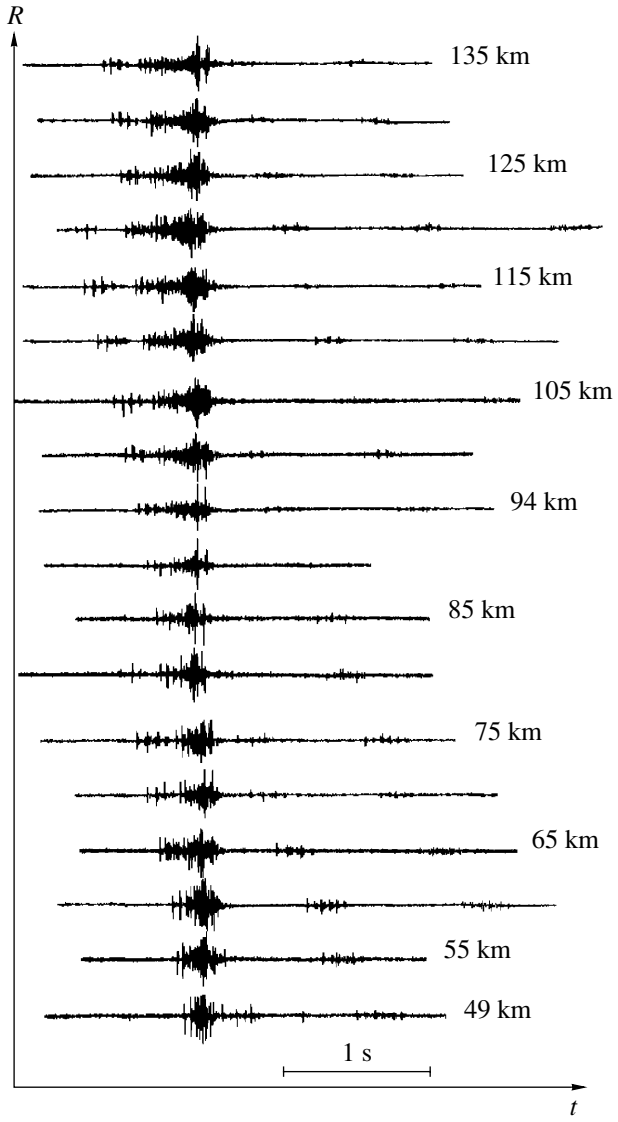


Fig. 6. Changes in the time structure of the sound field propagating in the USC of the western part of the Mediterranean Sea. The distance from the source varies from ~50 to ~135 km with steps of ~5 km. Each signal is normalized to its maximal amplitude.

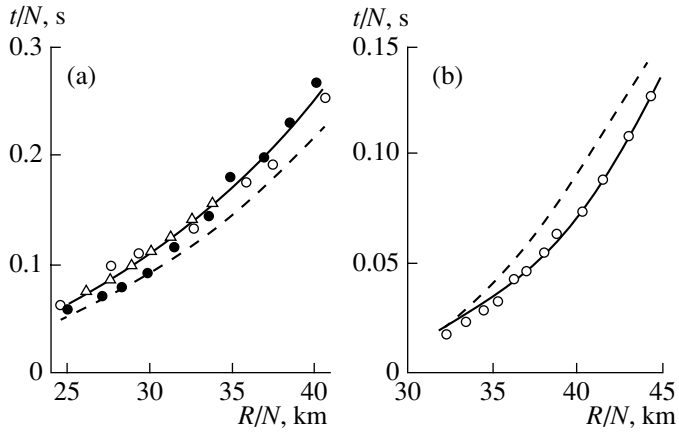


Fig. 7. (a) Reduced t/N–R/N diagram obtained from the experimental data for the western part of the Mediterranean Sea: experiment for N = (○) 2, (●) 3, and (△) 4; (—) approximation of experimental data by the dependence $t/N = (0.0155R/N)^{2.9}$; and (– – – –) the power law proposed in [5]. (b) Same for the eastern part of the Mediterranean Sea: (○ ○ ○) experiment (N = 4); (—) approximation of experimental data by the dependence $t/N = (0.0155R/N)^{5.6}$; and (— —) results of calculation.

sents the explosion-generated signals that were received at distances from 120 to 180 km with approximately 5-km steps in distance. As the distance increases, the classical quartets of time-separated elementary (single-path) signals become distinguished from the main group of the near-axis signals and the classical quartets are ahead of the main group. With a further increase in distance, the advance time of the quartets progressively increases and, finally, the quartet consists of surface reflected signals rather than purely water ones. At distances that are multiples of 42–45 km, the quartet vanishes because of bottom reflections. A new quartet arises that consists of signals differing by one lower ray half-cycle from the previous quartet. At a distance of 210 km from the source, no more than one isolated quartet and one "nascent" quartet (not fully separated from the main signal group) can be simultaneously observed in the total signal. Starting from a distance of 215 km and up to 360 km, two isolated quartets and one nascent quartet were observed, and so on.

Figure 7b shows the t/N–R/N curve plotted using the experimental data for a quartet branch corresponding to N = 4 and R = 120–200 km (the presentation of the t/N–R/N curves for both western and eastern parts of the Mediterranean Sea simplifies the comparison). This curve is well approximated by a power law of the form $t/N = (\alpha R/N)^p$ with the exponent $p = 5.6$ and the proportionality factor $\alpha = 0.0155$ (where t and R are expressed in seconds and kilometers, respectively). In the same figure, the results of computations with the Tebyakin [10] code are shown by the dashed curve. The calculated t/N–R/N curve passes somewhat higher than the experimental one. The computation used the profile $c(z)$ measured at the reception point. The sound speed at each horizon was recalculated from the bathymetric data with the use of the Wilson formula [14].

It should be noted that the time structures of the sound field are substantially different for the western and eastern parts of the Mediterranean Sea. The duration of the multipath explosion-generated signal (apart from bottom reflections) is more than two times higher in the western region than in the eastern one. The t/N–R/N curves are also different for these two parts of the sea. If one approximates the two curves by the function $t/N = (\alpha R/N)^p$ with the same α factor (0.0155), the values of the exponent p will be noticeably different (2.9 and 5.6).

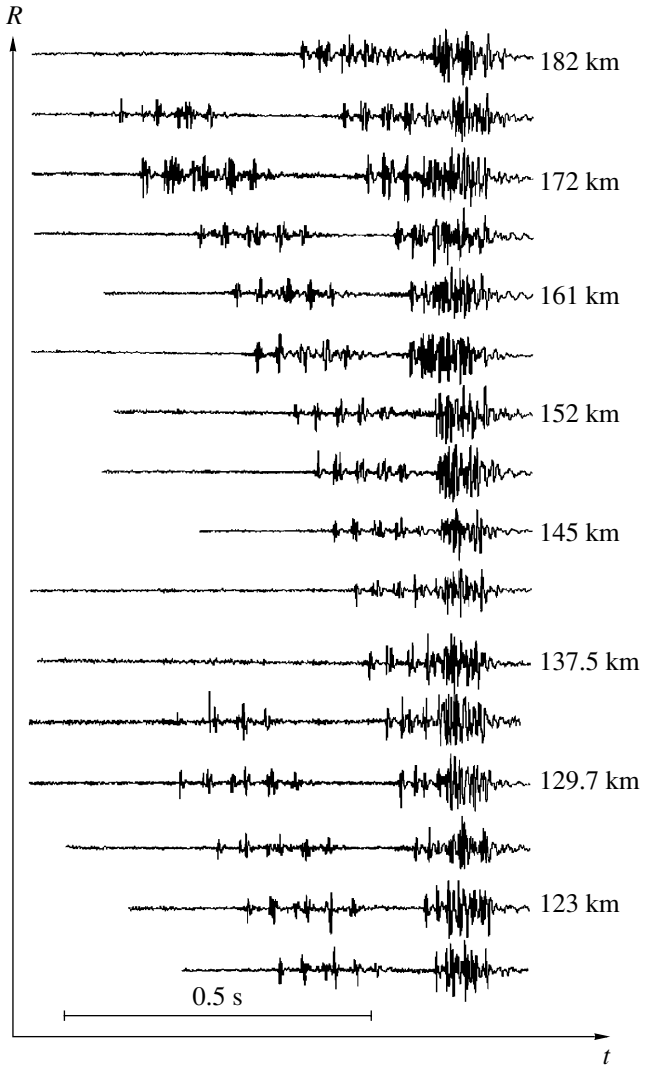


Fig. 8. Changes in the time structure of an explosion-generated signal propagating in the USC of the eastern part of the Mediterranean Sea. The distance from the source varies from ~120 to ~180 km with steps of ~5 km. Each signal is normalized to its maximal amplitude.

By frequency filtering of the explosion-generated signals, the geometric dispersion of the propagation velocity was estimated for the Mediterranean USC. The signal received at some distance from the source was filtered in 1/3-octave frequency bands covering 50–400 Hz. Time structures thoroughly matched in time were compared in different frequency bands for different distances from the source.

No noticeable dispersion was found for any signals received on the first path (if the dispersion exists in this frequency band, it is rather weak). For the signals received on the second path, the dispersion manifested itself at frequencies lower than 120–160 Hz. Low-frequency signals (50–60 Hz) propagate with appreciably lower velocity than high-frequency ones (125–160 Hz and higher). A similar situation was earlier observed by us in the Sea of Japan. This anomaly of the geometric dispersion in the USC was explained in [13].

One of the main objectives of the experiments on long-range propagation of explosion-generated signals consisted in studying the frequency dependence of sound attenuation in the Mediterranean Sea. Generally, the attenuation coefficient in the ocean is determined from the deviation of the experimental decay of the sound field level in the USC from the cylindrical spread law (strictly speaking, this method is valid only for channel-type propagation in a horizontally stratified medium). For explosion-generated signals, the following quantity is used as the sound field characteristic equivalent to the signal energy within a frequency band Δf:

$$E_f = \int_0^T p_f^2(t)dt,$$

where T is the duration of the explosion-generated signal and $p_f(t)$ is the acoustic pressure normalized to the frequency band Δf. The explosion-generated signals received and recorded in the experiment were computer-processed to obtain their power spectra. The bottom-reflected signals in the quartets are substantially lower than the near-axis signals of the main group in their energy. In the signal processing, these quartets were artificially rejected, and their contribution to the total sound field in the USC was neglected.

As a result of signal processing for both experiments, the decays of the sound field level were obtained, and they were different for different frequencies. The attenuation coefficients were determined from the deviation of the experimental decays from the cylindrical law on the path fractions within 50–640 and 120–610 km for the first and second experiments, respectively. At frequencies higher than 500 Hz, the far boundary of the path fraction used in estimating the attenuation was chosen to provide a sufficiently high signal-to-noise ratio (no less than 8–10 dB). Table 2 summarizes the attenuation coefficients obtained from the experiments with explosion-generated sound signals. The data presented correspond to the frequencies within 200–1000 Hz.

In addition to our data for the eastern part of the Mediterranean Sea, the data of Mellen *et al.* [6] are shown in Table 2. This experiment on long-range propagation of explosion-generated sound signals was performed in summer (July, 1982), in the Ionic Sea, on a 600-km path between the Sicily and Crete Islands.

Thus, Table 2 presents the data on low-frequency sound attenuation obtained on three different 600-km paths. The table also contains some results of estimating the attenuation coefficient with the use of noise-like 1/3-octave signals; these data are labeled with an asterisk.

The last column of Table 2 presents the absorption coefficients calculated according to the formulas of [15], where the temperature, salinity, and pH value

Table 2. Frequency dependence of the sound attenuation coefficient in the Mediterranean Sea (results of an analysis of experimental data)

Frequency, Hz	Attenuation coefficient, dB/km			Absorption coefficient, dB/km (calculations according to [15])
	western part of the sea	eastern part of the sea	eastern part of the sea [6]	
200	0.013	0.011	0.009	0.0054
250	0.016	0.015	0.013	0.0084
315	0.021	0.021	0.017	0.013
400	0.026	0.026	0.023	0.020
500	0.0335	0.038	0.034	0.030
630	0.042	0.044	0.052	0.044
800	0.050	0.058	0.073	0.064
1000	0.060	0.076, 0.08*	0.1	0.087
1250	–	–	0.13	0.113
1600	0.13*	0.09*	0.17	0.145
2000	–	0.2*	0.22	0.177
2500	–	0.21*	0.3	0.211
3150	–	0.235*, 0.240*	0.29	0.255
4000	0.320*	–	–	0.315
5000	–	0.4*	–	0.395

were specified to be 13°C, 38.5‰, and 8.1, respectively.

One can easily notice that the experimental attenuation coefficients obtained in the Mediterranean Sea at frequencies higher than 400 Hz agree well with the calculated absorption coefficients. Such an agreement is quite natural, because the formula given in [15] for the low-frequency absorption was, in particular, derived on the basis of our attenuation data.

The experimental attenuation coefficients also agree well with each other, although they were obtained in different parts of the Mediterranean Sea on different paths. In this sense, the attenuation coefficient in the Mediterranean Sea is a rather conservative characteristic, which is nearly independent of the water parameters substantially differing on the two paths. Such an independence of the experimental attenuation data on the hydrological environment implies that, in the case at hand, the only cause of attenuation is the sound absorption in sea water.

To conclude, let us summarize the main results of the long-range propagation experiments carried out in the Mediterranean Sea in different years.

(i) In the sound propagation conditions, a substantial difference exists between the western and eastern parts of the Mediterranean Sea. This difference is caused by the isolation of the two regions, by the weak water interchange between them and with the Atlantic Ocean, and by their different distance from the Strait of Gibraltar. The western region has an underwater sound channel with its axis at a depth of 100–200 m. In the eastern region, the USC axis is weakly pronounced and lies at a depth of 150–500 m (within the layer with depths from 200 to 400 m, the sound speed varies by no more than ±0.5–0.7 m/s).

(ii) The time structures of the sound field are also quite different for the western and eastern parts of the sea. The duration of the sound signal changes with a proportionality factor of 0.006 s/km in the western part, while this factor is 0.0026 s/km in the eastern region. The western and eastern parts of the sea are also different in their reduced $t/N–R/N$ diagrams: if they are described by a power-law function $t/N = (\alpha R/N)^p$, the exponent takes the values of 2.9 and 5.6 for the western and eastern regions, respectively, while the value of α remains the same (0.0155).

(iii) The zone structure of the sound field manifests itself for low (10–15 m) reception depths. The distance to the nearest boundary of the first convergence zone is 28.4 km for the western region and 33.2 km for the eastern region. A comparison between experiment and calculations shows a noticeable (up to 1.5 km) discrepancy in the positions of the first convergence zone. The most probable reason for such a discrepancy is the imperfection of the relations used to recalculate the salinity, water temperature, and hydrostatic pressure to the sound speed.

(iv) In spite of the substantial difference in the long-range propagation conditions, no noticeable distinctions between the western and eastern parts of the Mediterranean Sea were observed in sound attenuation and its frequency dependence. The experimental values obtained for attenuation coefficients on two 600-km

paths agree well with each other and with the data published by other researchers.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project nos. 01-02-16636 and 01-05-64711.

REFERENCES

1. V. F. Sukhoveĭ, *Seas of the World Ocean* (Gidometeoizdat, Leningrad, 1986).
2. B. S. Zalogin and A. N. Kosarev, *Seas* (Mysl', Moscow, 1999).
3. W. B. F. Ryan, E. Olausson, and R. W. Fairbridge, in *The Encyclopedia of Oceanography*, Ed. by R. W. Fairbridge (Reinhold, New York, 1966; Gidrometeoizdat, Leningrad, 1974).
4. *Hydrology of the Mediterranean Sea*, Ed. by I. M. Ovchinnikov, E. A. Plakhin, *et al.* (Gidrometeoizdat, Leningrad, 1976).
5. C. C. Leroy, in *Underwater Acoustics*, Ed. by V. M. Albers (Plenum, New York, 1967; Mir, Moscow, 1970), Vol. 2.
6. R. H. Mellen, T. Akal, E. H. Hug, and D. G. Browning, J. Acoust. Soc. Am. **78**, S70 (1985).
7. D. Yu. Mikhin, O. A. Godin, Yu. A. Chepurin, *et al.*, in *Acoustics of the Ocean: Proceedings of the Advanced-Study School of Academician L. M. Brekhovskikh* (GEOS, Moscow, 1998), pp. 24–30.
8. R. A. Vadov, Akust. Zh. **41**, 202 (1995) [Acoust. Phys. **41**, 172 (1995)].
9. R. A. Vadov, Akust. Zh. **49**, 278 (2003) [Acoust. Phys. **49**, 230 (2003)].
10. V. P. Tebyakin *et al.*, *Raymod52—A Software for Calculating the Sound Fields in a Layered Inhomogeneous Ocean in Terms of Ray Acoustics: A Report* (Inst. of Acoustics, USSR Acad. Sci., Moscow, 1990).
11. R. A. Vadov, in *Proceedings of 13th Session of the Russian Acoustical Society* (GEOS, Moscow, 2003), Vol. 4, pp. 27–32.
12. M. Greenspan and C. E. Tschiegg, in *Underwater Acoustics*, Ed. by V. M. Albers (Plenum, New York, 1963; Mir, Moscow, 1965).
13. R. A. Vadov, Okeanologiya (Moscow) **40** (2), 165 (2000).
14. W. D. Wilson, J. Acoust. Soc. Am. **32**, 1357 (1960).
15. R. A. Vadov, Akust. Zh. **46**, 624 (2000) [Acoust. Phys. **46**, 544 (2000)].

Translated by E. Kopyl

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 265–272. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 322–330.

Methods of Estimating the Acoustic Parameters of Ground

S. A. Rybak*, S. A. Makhortykh, S. A. Kostarev***, and A. V. Derguzov****

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

e-mail: rybak@akin.ru

*** Institute of Mathematical Problems of Biology, Russian Academy of Sciences, ul. Institutskaya 4, Pushchino, Moscow oblast, 142290 Russia*

e-mail: makh@impb.ru

**** Tunnel Association of Russia, ul. Sadovo-Spasskaya 21, Moscow, 107217 Russia*

e-mail: viblab_tar@spacenet.ru

Received September 4, 2003

Abstract—A method for determining the dynamic and dissipative characteristics of the ground (velocities and attenuation coefficients of elastic waves) in urban environments is presented. The knowledge of these characteristics is necessary for estimating and monitoring environmental conditions in the presence of intense sources of sound and vibration. Such problems arise in connection with prediction of vibration levels produced on the ground surface by underground and surface sources (transportation, construction machinery, etc.). With allowance for the specific features of solving direct problems of elastic wave field calculation, inverse problems of the tomography of ground are formulated and methods of their solution are described. Several algorithmic realizations based on multiply solving the direct problem and also on expanding the total field in generalized spectral components are considered. The latter approach makes it possible to simplify the estimation of parameters by applying the minimization of the residual functional. The basis functions used for approximating the experimental data are chosen in the form of Jacobi polynomials. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The problems of vibroacoustic monitoring in urban environments are rather topical because of the multitude of active sources, which include subway lines, railroads, construction machinery, etc. Ground is a fairly complex geophysical medium characterized by inhomogeneity and complicated absorbing and elastic properties. From the mathematical point of view, a simplified model of a homogeneous medium with a free surface, shear and bulk elasticities, and dissipation is already difficult to analyze. Computational difficulties sometimes lead to the loss of essential components of wave solutions (specifically, in the case of an incorrect application of finite-difference schemes). If the ground consists of many layers with different elastic and mass parameters, the pattern of the vibration field becomes much more complicated. In addition to mathematical difficulties, actual computations are hindered by the partial or total absence of data on the geometric parameters and elastic properties of ground at a specific site of interest.

Because of the complex structure of the vibration field and the lack of initial data, the methods commonly used in practice are mainly empirical. The behavior of the error caused by an inaccurate presetting of the acoustic parameters of the medium is shown in Fig. 1. The velocities of longitudinal and transverse waves are assumed to be preset with errors of 100 and 20 m/s, respectively, and their mean values are 600 and 200 m/s. The dependences shown in Fig. 1 are obtained for 31.5- and 63-Hz octave frequency bands.

Direct problems of vibroacoustic prediction are described in [1]. In developing the methods for solving direct and inverse problems (acoustic prediction and estimation of the parameters of ground, respectively), one should take into account that the measurement accuracy in actual conditions is about $\Delta L = 20\log\frac{\Delta a}{a} \sim$ 2 dB, where Δa is the measurement accuracy and a is the absolute value of a given vibration characteristic (displacement, velocity, or acceleration). Therefore, a natural requirement for the computational models (and, hence, for the accuracy of estimates of the parameter values) is that they provide the accuracy given above.

According to the models [1] proposed for calculating the vibration field in an elastic medium, it is necessary to preset the velocities of longitudinal and transverse waves (or pairs of parameters: Young's modulus and Poisson's ratio, Lamé constants λ and μ) and the density. In a stratified medium, the values of the parameters are preset for each layer. For the inverse problem, the parameters listed above are the desired quantities.

The dynamic parameters of ground that are necessary for predicting the vibration amplitudes or their log-

1063-7710/04/5003-0265$26.00 © 2004 MAIK "Nauka/Interperiodica"

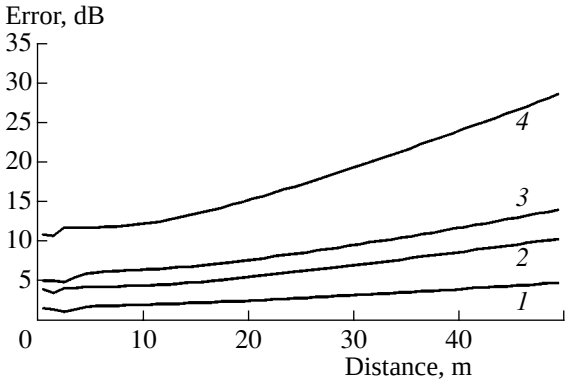


Fig. 1. Dependence of the error of determining the elastic wave field amplitude on the distance from the source when the acoustic parameters of the medium are inaccurately preset: the horizontal displacement component at frequencies of (*1*) 31.5 and (*2*) 63 Hz and the vertical component (*3*, *4*) at the same frequencies. The velocities of longitudinal and transverse waves are preset with errors of 100 and 20 m/s, respectively, and their average values are 600 and 200 m/s, respectively.

arithmic levels are usually determined in practice from direct geological measurements or by using tabular values. The latter approach is often unacceptable because of the wide scatter of parameters for each type of ground. Section 1 of this paper is devoted to the general statement of the problem of estimating the given parameters and describes the method of solving this problem by multiply solving the direct problem, according to the approach proposed in [1]. The sequence of steps for estimating the parameters from the measured vibration characteristics and *a priori* geological data is described. Section 2 presents an example of applying the proposed procedure. It also presents a general approach to describing acoustic data by the generalized spectral-analytical method [2], which is used in the processing of experimental data and in solving the problem of estimating the parameters of ground.

1. PROCEDURE OF ESTIMATING THE ELASTIC, MASS, AND DISSIPATIVE PARAMETERS OF GROUND

According to the general statement of the problem and the requirements of the procedure for calculating the vibration characteristics on the ground surface, it is necessary to determine the following parameters: the type of stratification, i.e., the number of layers and the thickness of each layer (below, we consider a total number of layers $1 \le N_c \le 3$ with noticeably different properties); the density of ground ρ; the velocities of longitudinal and transverse elastic waves c_l and c_t; and the attenuation coefficient β in each layer. The maximal number of layers is taken to be equal to 3 with allowance for the calculations performed for real urban environments. This number is usually sufficient for obtaining the required accuracy in estimating vibration accelerations of 2–3 dB.

If the dynamic and dissipative properties of two adjacent layers differ by a factor of less than 1.5, these layers can be combined into one with a total thickness of $h = h_1 + h_2$ and with the average propagation velocities and attenuation coefficient of elastic waves:

$$c_{l,t} = \frac{c^1_{l,t}h_1 + c^2_{l,t}h_2}{h_1 + h_2}, \quad \beta = \frac{\beta_1 h_1 + \beta_2 h_2}{h_1 + h_2}. \qquad (1)$$

Here, subscripts 1 and 2 refer to the layer numbers and subscripts l and t refer to longitudinal and transverse waves. It is necessary to consider only the upper part of ground, to a depth of

$$H = h_s + 5 \text{ (m)}, \qquad (2)$$

where h_s is the distance from the ground surface to the source.

In determining the structure of the upper part of ground in the preliminary analysis of the geological situation, one should use available geological data. On the basis of *a priori* information and the aforementioned rule of combining layers with close properties, the initial one-, two-, or three-layer model of ground is constructed. The preset values for the layer thicknesses are ultimate, while the dynamic and dissipative parameters of the layers require further refinement. If preliminary information on the geological structure of ground is absent, it is necessary to solve a complete inverse problem with an unknown number of layers and their unknown thicknesses and dynamic and dissipative properties. Since, in urban environments, the density of ground varies only slightly (1600–2000 kg/m^3), this variation can be neglected within acceptable accuracy and the density can be taken to be equal to ρ = 1800 kg/m^3. Poisson's ratio of ground in urban areas varies from 0.1 to 0.45. However, estimates show that the specific value of this ratio little affects the vibration characteristics. Therefore, within acceptable accuracy, Poisson's ratio can also be considered as constant and taken to be equal to its average value $\nu = 0.4$.

The solution of the inverse problem of estimating the parameters of ground is performed by the trial-and-error method, namely, by substituting their specific values into the computational scheme and comparing the results of calculations with the results of measurements at different distances from the source of vibration. The procedure should be as follows. The lowest values of longitudinal wave velocity and the minimal values of attenuation coefficient for a given ground are taken as the initial approximation. The maximal values of the same parameters for the corresponding ground are also preset. With allowance for the constancy of Poisson's ratio, the transverse wave velocity in the ground is determined by the formula

$$c_t = c_l\left[\frac{1-2\nu}{2(1-\nu)}\right]^{1/2}. \qquad (3)$$

If the required agreement with experiment is not achieved in the first run, it is necessary to vary the values of the parameters under determination:

$$\beta = \beta_{\min} + i\Delta\beta, \quad i = 1, 2, 3 \tag{4}$$

at a step of

$$\Delta\beta = (\beta_{\max} - \beta_{\min})/3, \tag{5}$$

$$c_j = c_{\min} d^j, \quad j = 1, 2, 3, \ldots . \tag{6}$$

Here, $\beta_{\max}$, $c_{\max}$ and $\beta_{\min}$, $c_{\min}$ are the maximal and minimal values of the attenuation coefficient and longitudinal wave velocity, respectively, for the given type of ground. The coefficient d is chosen depending on the ratio $c_{\max}/c_{\min}$, for example, by the rule given in Table 1.

Thus, depending on the value of $c_{\max}/c_{\min}$, we have three, four, or five iterations in longitudinal wave velocity. The transverse wave velocity is calculated at each iteration by formula (3). If necessary, the attenuation coefficient is varied at each iteration. When the required accuracy is achieved at some step, the problem of estimating the parameters is solved. If no data on the geological structure of ground in the region under study are available, the information on the stratification is eliminated from the initial data. Then the number of layers (from 1 to 3) and their thicknesses become additional unknown parameters. The analysis begins with the simplest one-layer model. If, with this model, the required accuracy of fitting with the experimental data is not achieved, one should pass to two- and three-layer models. The initial values of the longitudinal wave velocity and attenuation coefficient are chosen to be $c = 100$ m/s and $\beta = 0.05$.

The procedure described above was automated using the VibraCalc program developed by the authors of this paper. The initial data for calculations is a set of experimental values of vibration characteristics measured at given distances from a source of standard intensity.

The proposed approach allows one to obtain adequate estimates of parameters if the properties of ground vary slowly in the horizontal direction. For example, for a wave frequency of 30 Hz, the scale of variations should exceed 40 m for sandy ground (i.e., about two wavelengths). The resulting estimates of parameters can be used for calculating the wave field in direct problems. Otherwise, the problem should be solved in a full three-dimensional formulation.

2. RESULTS OF USING A SIMPLIFIED SCHEME OF PARAMETER ESTIMATION

The approach described above was tested using experimental data obtained from the Tunnel Association of Russia. The data included vibration characteristics on the ground surface. The wave field was excited by a series of weak standardized explosions in a 22-m-deep well. Vertical and horizontal components of vibration acceleration were measured at distances $l = 0$, 10, 20, 30, and 40 m from the well. The results of measurements are presented in Fig. 2 (Figs. 2a and 2b refer to vertical and horizontal vibration accelerations, respectively).

Table 1

$c_{\max}/c_{\min}$	d
≤3	$\sqrt{c_{\max}/c_{\min}}$
≤6	$\sqrt[3]{c_{\max}/c_{\min}}$
>6	$\sqrt[4]{c_{\max}/c_{\min}}$

In the case under consideration, preliminary information on the stratification of the upper part of ground was available. The initial parameters used for the calculation were the values shown in Table 2. The density of ground and Poisson's ratio were $\rho = 1800$ kg/m^3 and $\nu = 0.4$.

For the given stratification, the criterion of a correct determination of the parameters of ground is the minimum of the functional

$$F(m) = \sqrt{\sum_{i=0}^{N} W(x_i)\left(f(x_i, m_j) - \phi(x_i)\right)^2}. \tag{7}$$

Here, $\phi(x_i)$ represents the experimental values of the vibration amplitude at the points of measurement x_i, $f(x_i, m)$ represents the calculated values of vibration amplitude at the same points x_i, and $W(x_i)$ is the weighting function taking into account the reliability measurements. Calculations were performed using the function shown in Fig. 3. The quantities m_j are actually the parameters to be determined: $m_1 = c_1$, $m_2 = \beta_1$, $m_3 = c_2$, and $m_4 = \beta_2$ (the subscripts of c and β refer to two different layers). The results of estimating the parameters are presented in Table 3.

The comparison of the calculated (with the values of parameters that were determined) and experimental data (the levels L of the vertical projections of vibration acceleration for 16-, 31.5-, and 63-Hz octave bands) are shown in Fig. 4. The levels were calculated by the formula $L = 20\log(a/a_0)$ with a threshold value of $a_0 = 10^{-6}$ m/s^2.

The results of the calculations agree well with the measurements at small distances, and the accuracy of these calculations increases with growing frequency. This result testifies to the optimal choice of the weighting function for the 63-Hz octave band and to the need to correct it for frequencies of 16 and 31.5 Hz. Another way of improving the accuracy of estimates can be the choice of a more detailed stratification model, for example, a three-layer model of ground. The comparison of the results of calculations with the measurements

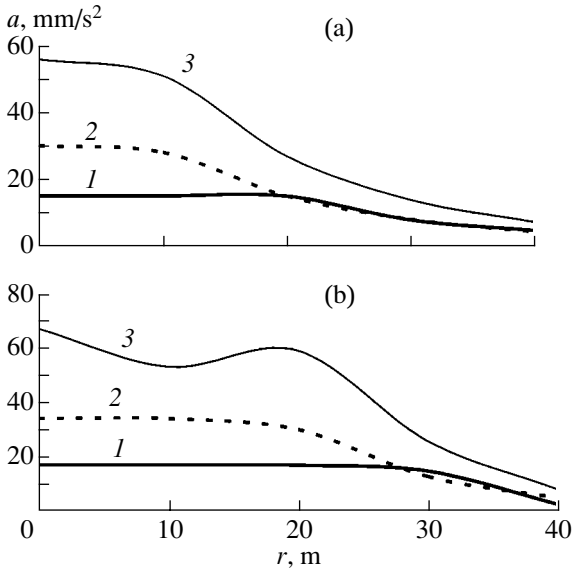


Fig. 2. Results of measuring the wave field amplitudes produced by a series of weak standardized explosions in a well: dependences of the (a) vertical and (b) horizontal components of vibration acceleration on distance for frequencies of (*1*) 16, (*2*) 31.5, and (*3*) 63 Hz.

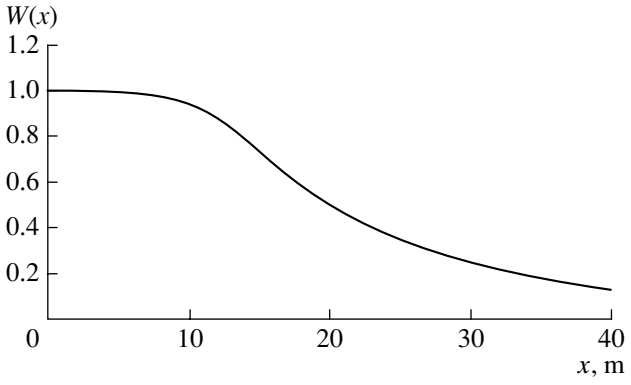


Fig. 3. Dependence of weighting function (7) on distance.

for the ground model specified in Table 4 is illustrated in Fig. 5.

3. ORTHOGONAL EXPANSIONS

In this section, we suggest an approach to describing acoustic data and to a parametric identification of systems on the basis of orthogonal expansions. The method uses an adaptive tuning of the approximation of a digital array by a truncated orthogonal series. All subsequent procedures of the analysis are performed in the expansion coefficient space. This approach considerably reduces the computation time and allows an efficient use of analytical methods in the problems of data processing.

The essence of the method is as follows. As the data are supplied, an optimal set of functions is automatically chosen for each of the signals with the aim to obtain the spectral expansion in terms of these functions. To achieve a given accuracy (in the root-mean-square or uniform sense) of the analytical description, the set of functions is adapted to the properties of a given signal, which leads to an efficient compression of the volume of data and to the approximation of the signal by an expression of minimal complexity.

An analytical description of data by fragments of orthogonal series has a constant structure. Therefore, one can easily obtain analytical dependences for finding different estimates and characteristics. As the basis functions, one can use regular trigonometric Fourier series or a wider set of functions. Below, we use classical orthogonal polynomials [2]. The latter have advantages due to the multitude of sets of basis functions, the presence of parameters that allow an easy modification of the description and its adjustment to specific signals, and the possibility of obtaining exact analytical dependences in the expansion coefficient space. An explicit use of adaptive procedures makes it possible to optimize the description and to obtain the least complex expressions with the required accuracy of approximation.

Systems of basis functions are complete and closed. Therefore, any function $x(t)$ belonging to space L^2 can be exactly represented by an infinite orthogonal series of any orthonormal classical polynomials or functions:

$$x(t) = \sum_{n=0}^{\infty} A_n \varphi_n(t), \tag{8}$$

where $\varphi_n(t)$ are orthonormal functions from the chosen basis and A_n are the expansion coefficients of $x(t)$ expanded in the chosen basis. The coefficients A_n are calculated by the well-known formula

$$A_n = \int_0^T x(t)\varphi(t)\rho(t)dt, \tag{9}$$

where $\rho(t)$ is a weighting function.

In actual conditions, the number of terms in the series is finite and the problem consists in obtaining a given accuracy of approximation by the least complex expression. To satisfy the condition $N = N_{\min}$, it is necessary to chose a basis whose first functions have the form close to or coincident with the form of the signal under study. To obtain an integral estimate of the form of the input signal, we introduce the "form factor":

$$K_{\text{TM}} = \int_0^T x(t)\eta(t)dt \Big/ \int_0^T x(t)\xi(t)dt, \tag{10}$$

where $x(t)$ is the signal to be approximated and $\eta(t)$ and $\xi(t)$ are known functions chosen depending on which

Table 2

Layer no.	Type of ground	Layer thickness, m	c_{min}, m/s	c_{max}, m/s	β_{min}	β_{max}
1	Clay soil	6	500	2800	0.1	0.4
2	Dry sandy soil	20	150	900	0.05	0.1

Table 3

Layer no.	Longitudinal wave velocity c, m/s			Attenuation coefficient β		
	16 Hz	31.5 Hz	63 Hz	16 Hz	31.5 Hz	63 Hz
1	800	900	1000	0.2	0.15	0.1
2	600	700	800	0.15	0.1	0.06

general properties of signal $x(t)$ should be taken into account. The values of the form factor obtained for the pair $\eta(t)$, $\xi(t)$ can be calibrated and related to one or another basis; it can then be used in calculations for determining the optimal set of basis functions.

In modeling the wave field and solving the parametric identification problem in spectral formulation, we proceed from the expression

$$y(x) = M(a(x), \alpha_i(x)), \tag{11}$$

where $y(x)$ and $a(x)$ are the input and output signals, $\alpha_i(x)$ is the set of varied parameters, and M is a given

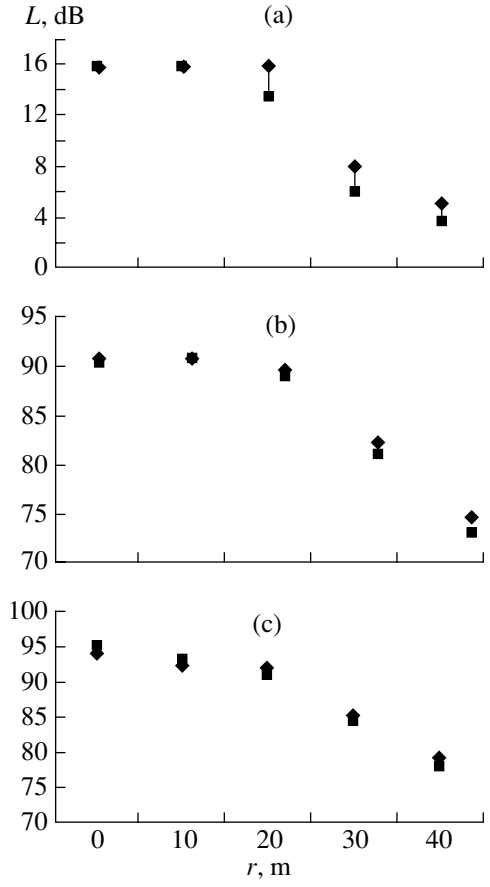


Fig. 4. Comparison of calculated (full squares) and experimental (full diamonds) data (levels L of the vertical projections of vibration acceleration) in different octave bands: (a) 16, (b) 31.5, and (c) 63 Hz.

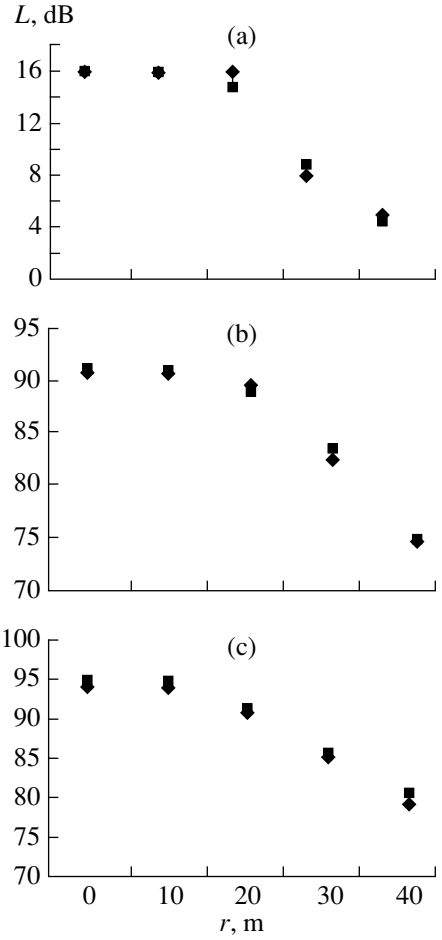


Fig. 5. Comparison of the results of calculations (full squares) with measurements (full diamonds) in different octave bands for the ground model specified in Table 4: (a) 16, (b) 31.5, and (c) 63 Hz.

Table 4

Layer no.	Layer thickness, m	Longitudinal wave velocity c, m/s			Attenuation coefficient β		
		16 Hz	31.5 Hz	63 Hz	16 Hz	31.5 Hz	63 Hz
1	6	750	900	1000	0.2	0.15	0.1
2	6	600	700	800	0.15	0.1	0.06
3	14	1200	1250	1250	0.15	0.1	0.06

operator. If the input signal and the parameters are known, we have a direct problem. Otherwise, we obtain an inverse problem of estimating the parameters $\alpha_i(x)$ of the signal $a(x)$ from the known experimental data $y(x)$. In some cases, the solution of the inverse problem can be reduced to the solution of the direct problem by multiply solving the latter (the method used above).

Solution of the direct problem with the use of orthogonal expansions is illustrated below by a simple example. Let us consider a system described by a one-dimensional (depending on a single independent variable x) linear differential operator

$$\hat{L} = \frac{d^k}{dx^k} + a_1\frac{d^{k-1}}{dx^{k-1}} + \ldots + a_{k-1}\frac{d}{dx} + a_k: \qquad \hat{L}y = f \tag{12}$$

with a boundary condition at the ends of interval s (at points x_0 and x_1) in the form

$$\hat{g}y|_s = f_0, \tag{13}$$

where $\hat{g} = \dfrac{d^{k-1}}{dx^{k-1}} + b_1\dfrac{d^{k-2}}{dx^{k-2}} + \ldots + b_{k-2}\dfrac{d}{dx} + b_{k-1}$.

Here, $y(x)$ is the desired function of the single variable. For simplicity, we assume that the quantities a_i and b_i are constants. The general structure of the algorithm of solving the above-stated problem is as follows. We write in the general form the expansions of the known and desired functions in the chosen orthogonal basis, for example, in the Chebyshev polynomials of the first kind $\{T_i(x)\}$:

$$y = \sum_{i=0}^{N} A_i T_i, \quad f = \sum_{i=0}^{N} B_i T_i \text{ and } \sum_{i=0}^{N} A_i \hat{g} T_i|_s = f_0 \tag{14}$$

(for solution $y(x)$ and all the known functions from Eqs. (12) and (13)). We substitute expansions (14) into Eqs. (12) and (13) and then use the known analytical expressions (e.g., the change from the expansion coefficients of functions to expansion coefficients of derivatives, products, etc.). As a result, we obtain an algebraic system of linear or nonlinear equations in unknown expansion coefficients A_i with known coefficients B_i. The solution of this system yields approximation (14) of the solution to problem (12), (13). The smoothness of the solution is provided by the properties of basis functions, and the required accuracy is achieved by the appropriate choice of the series expansion length.

The derivative of the signal (function) is expanded in a Fourier series $y'_N(t) = \sum_{i=0}^{N} D_i^{(1)} T_i(t)$ on the condition that $\mathbf{D}^{(1)} = \mathbf{D}_N\mathbf{A}$. Here, $\mathbf{D}^{(1)}$ and $\mathbf{A}$ are N- and $(N+1)$-dimensional column vectors of the expansion coefficients of the derivative and the initial function (the superscript 1 refers to the first derivative of the function), and $\mathbf{D}_N$ is an $(N+1)\times N$ matrix representing the approximation of the differential operator in the chosen functional basis $\mathbf{D}_N = \{d_{ij}^{(1)}\}$. Continuing the procedure described above, we obtain an expression for the differential operator of arbitrary order k: $\{d_{ij}^{(k)}\}$ (an $(N+1)\times(N-k+1)$) matrix). Correspondingly, the vector of expansion coefficients of the kth derivative has the form $\mathbf{D}^{(k)} = \mathbf{D}_{N-k+1}\cdot\ldots\cdot\mathbf{D}_N\mathbf{A}$. In the case of a set of orthogonal Chebyshev polynomials of the first kind, the matrix of the derivative has the form

$$\mathbf{D}_N = \overbrace{\begin{Bmatrix} \left(\begin{array}{c|cccccc} & \sqrt{2} & 0 & 3\sqrt{2} & 0 & 5\sqrt{2} & \ldots \\ & 0 & 4 & 0 & 8 & 0 & \ldots \\ & 0 & 0 & 6 & 0 & 10 & \ldots \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & 0 & 0 & \ldots & 0 & 0 & 2N \end{array}\right) \end{Bmatrix}}^{N} N-1 \tag{}$$

Figure 6 shows the solution to the test problem (the wave field in a planar waveguide with specific boundary conditions):

$$u'' + 3u = 0, \quad u(0) = 1, \quad u'(5) = 2\sqrt{3}\cos(5\sqrt{3}) - \sqrt{3}\sin(5\sqrt{3}). \tag{15}$$

The approximate solution in the form of a truncated series expansion in Chebyshev polynomials ($N = 8$, curve *2*) is compared with the known analytical solution (curve *1*) $u(x) = \cos(\sqrt{3}x) + \sin(\sqrt{3}x)$. Figures 6b and 6c show the error curves for $N = 8$ and 16, respectively.

A procedure analogous to that proposed for solving the direct problem can be used for obtaining the expansion of a function describing an unknown parameter (e.g., the distribution of the elastic wave velocity in depth), i.e., for solving the inverse problem. In this case, the Fourier coefficients for the input and output signals are preset and the desired spectrum for $\alpha_i(x)$ is calculated using the model approximation given by Eq. (12). If a direct calculation is difficult, an iteration procedure can be used. A sufficiently low parametric complexity of the problem is provided by an appropriate choice of the approximation basis. The number of unknown parameters should correspond to the number of expansion terms.

The proposed method was used to solve the inverse problem for a waveguide in the ground. Such a waveguide is formed, for example, in the course of freezing or melting of the upper ground layer, which leads to its saturation with water. In both cases, the propagation velocity of elastic waves (primarily, longitudinal waves) near the surface is higher than in the depth.

The sound pressure p in soft ground (the fluid-medium approximation) is described by the equation

$$\frac{d^2p}{dz^2} + \omega^2\left[\frac{1}{c^2(z)} - \frac{1}{c_{\text{ph}}^2}\right]p = 0, \qquad (16)$$

and the velocity distribution $c(z)$ in the case under consideration is shown in Fig. 7a. Here, z is the vertical coordinate and c_{ph} is the phase velocity in the waveguide. The source intensity is adjusted to fit the measured values.

For approximating the solution to Eq. (16), we use modified Jacobi polynomials $R_n^{\alpha\beta}(x)$ [2] defined on the interval $(0, +\infty)$. This basis provides a good approximation of $p(z)$ by a series of length $N = 4$. The sound velocity distribution $c(z)$ obtained from the inversion of Eq. (16) is shown in Fig. 7b. Its approximation is performed by a single polynomial with number $n = 2$ and parameters $\alpha = 2$ and $\beta = 1$.

Often, in calculating the desired Fourier coefficients, the iteration procedure proves to be more efficient than the direct use of Eqs. (12)–(14). The main idea of this procedure consists in the use of the known initial approximation for coefficients $A_i^{(0)}$. Choosing the initial expansion $y_0(t) = \sum_{i=0}^{n} A_i^{(0)} T_i(t)$, one can obtain a refined expression $y_1(t) = \sum_{i=0}^{n} A_i^{(1)} T_i(t)$ if the initial model can be represented in the form $y(t) = F(y(t), t)$, where F is an explicitly defined function or operator.

The application of the simplified procedure described at the beginning of this paper allows one to obtain adequate estimates of the parameters of ground in most real situations with soft or moderately hard grounds whose properties vary slowly in the horizontal direction. The presence of *a priori* geological information is useful for obtaining more reliable results but is not necessary. Knowledge of the distribution of vibration characteristics over the ground surface allows one to obtain a good approximation of vibration in the upper layers of ground. The distributions of elastic and dissipative properties can also be obtained. To refine the results, one can use a more elaborate measuring procedure, which includes the determination of vibration characteristics not only on the surface but in the depth of the ground. Data obtained from such experiments can be used in a more accurate method based on the orthogonal expansions of records.

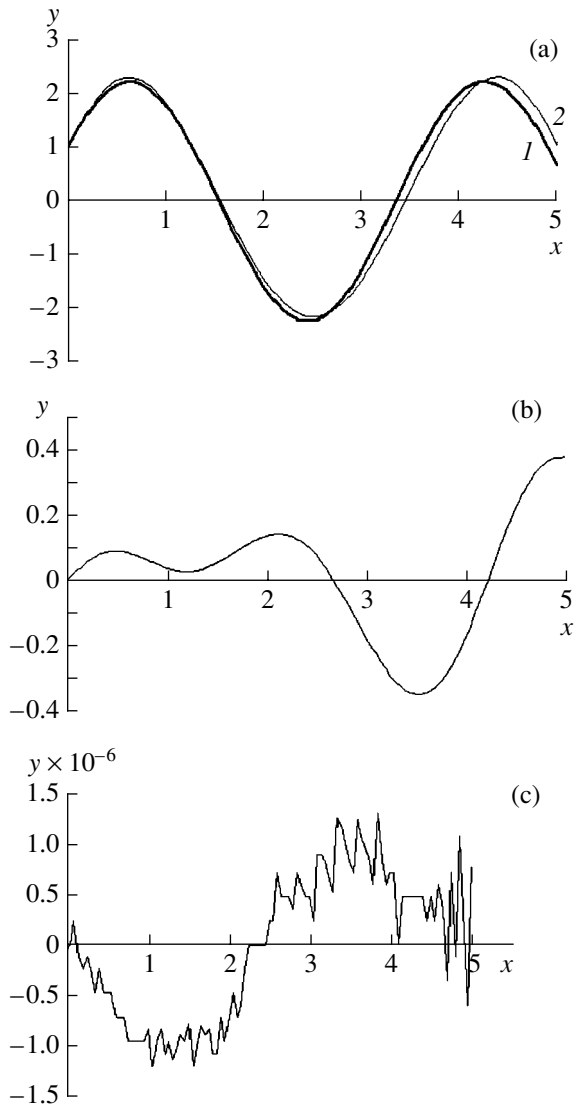


Fig. 6. Solution to test problem (15): (a) (*1*) the exact analytical solution and (*2*) the approximate solution in the form of a truncated series expansion in Chebyshev polynomials ($N = 8$). Error curves for $N =$ (b) 8 and (c) 16.

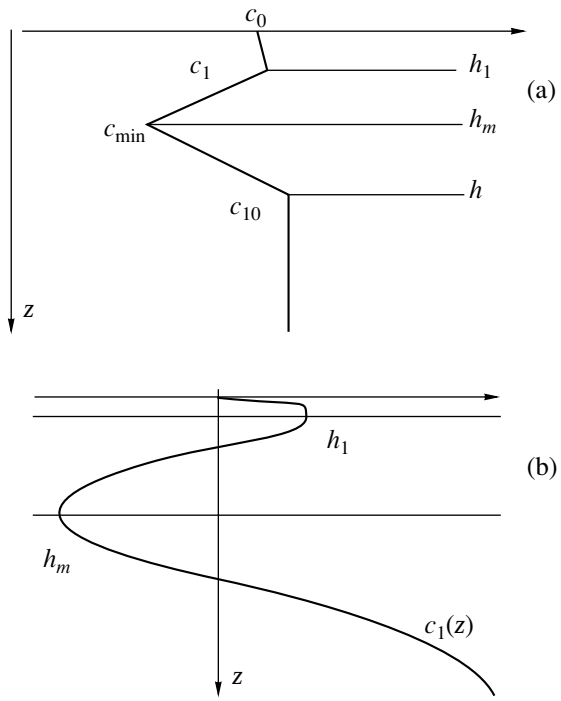


Fig. 7. (a) Sound velocity distribution $c(z)$ in Eq. (16). (b) Approximation of the solution to the inverse problem in the modified Jacobi polynomial basis.

The change from a set of exponents to orthogonal polynomial series in the representation of the behavior of the acoustic field provides a possibility to improve the accuracy of approximation to a considerable extent and to obtain more reliable estimates of the desired elastic and dissipative parameters.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 02-02-17143 and 01-02-16127) and the 6th Competition of Young Scientists of the Russian Academy of Sciences (project no. 107).

REFERENCES

1. S. A. Kostarev, S. A. Makhortykh, and S. A. Rybak, in *Noise and Vibration from High-Speed Trains* (Thomas Telford, London, 2001), pp. 397–422.
2. F. F. Dedus, S. A. Makhortykh, M. N. Ustinin, and A. F. Dedus, *Generalized Spectral-Analytical Method of Data Array Processing* (Mashinostroenie, Moscow, 1999).

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 273–277. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 331–336.

Helical Waves of an Elastic Cylindrical Shell

V. V. Tyutekin

Andreev Acoustics Institute, Russian Academy of Sciences, ul. Shvernika 4, Moscow, 117036 Russia
e-mail: Tyutekin@akin.ru

Received July 15, 2003

Abstract—Properties of helical waves of a cylindrical shell described by Kirchhoff–Love equations are considered. The problem is reduced to the case of the propagation of plane waves in an equivalent plate. On the basis of the corresponding dispersion equation and its solution, a conclusion is made about the anisotropy of the shell properties. Dispersion curves are plotted for different angles of propagation of helical waves with respect to the shell axis. Displacements of the shell along and across the direction of wave propagation are calculated. © *2004 MAIK "Nauka/Interperiodica".*

Problems of wave propagation in cylindrical shells, including those immersed in an acoustic medium, remain topical from the point of view of both theory and practical applications [1–7]. The most commonly used method of solving this kind of problem consists in describing an elastic cylindrical shell by equations based on the Kirchhoff–Love hypothesis [8]:

$$L\mathbf{U} = 0, \tag{1}$$

where $\mathbf{U} = (u, v, w)^T$ is the shell displacement vector and L is the matrix differential operator, which has the form

$$L_{11} = \frac{\partial^2}{\partial z^2} + \frac{1+\sigma}{2a^2}\frac{\partial^2}{\partial \theta^2} + k_p^2;$$

$$L_{12} = L_{21} = \frac{1+\sigma}{2a}\frac{\partial^2}{\partial z \partial \theta}; \quad L_{13} = -L_{31} = -\frac{\sigma}{a}\frac{\partial}{\partial z};$$

$$L_{22} = \frac{1-\sigma}{2}\frac{\partial^2}{\partial z^2} + \frac{1}{a^2}\frac{\partial^2}{\partial \theta^2} + k_p^2; \quad L_{23} = -L_{32} = -\frac{1}{a^2}\frac{\partial}{\partial \theta};$$

$$L_{33} = -\left(\frac{h^2}{12}\Delta^2 + \frac{1}{a^2} - k_p^2\right).$$

Here, z and θ are cylindrical coordinates, u and v are displacements along the z and θ axes, w is displacement perpendicular to the shell surface, a is the radius of the shell, and h is its thickness. In addition, $k_p^2 = \frac{\rho\omega^2(1-\sigma^2)}{E}$ is the square of the longitudinal wave number in a plate, ρ is density, E is Young's modulus, σ is Poisson's ratio, and $\Delta = \frac{\partial^2}{\partial z^2} + \frac{1}{a^2}\frac{\partial^2}{\partial \theta^2}$.

In solving the problems of wave propagation along the axis of a shell described by Eq. (1), one usually considers solutions that are periodic in angle θ. Their dependence on θ is described by $\cos n\theta$ or $\sin n\theta$, where $n = 0, 1, \ldots$; i.e., one obtains a standing wave in θ. This wave can be represented as two travelling waves $e^{in\theta}$ and $e^{-in\theta}$ (with the corresponding coefficients), and, hence, the total solution representable in the form $e^{i(k_z z \pm n\theta)}$ is a sum of two waves of a helical type (where n is an integer and k_z is the projection of the wave number on the z axis).

This paper considers the properties of helical waves of an elastic cylindrical shell with an arbitrary parameter ν. In this case, the solution can be represented in the form

$$\mathbf{U} = \mathbf{U}_0 e^{i(k_z z \pm \nu\theta)}, \tag{2}$$

where $\mathbf{U}_0 = (u_0, v_0, w_0)^T$ is the vector of displacement amplitudes. In the general case, formula (2) corresponds to aperiodic solutions characteristic of nonclosed shells or longitudinally fixed shells (including longitudinally cut ones). These solutions may assume that, in the general case, angle θ varies within $-\infty < \theta < \infty$; i.e., the solutions are considered on a many-sheeted Riemann surface.

Equation (1) allows one to replace a cylindrical shell that is infinite in two dimensions by an infinite flat plate. This can be done through the evident substitution $\nu\theta = k_y y$, where $y = a\theta$ is a Cartesian coordinate (the y axis is perpendicular to the z axis) and $k_y = \frac{\nu}{a}$ is the projection of the wave number k on the y axis. In the framework of this approach, helical waves (2) (Fig. 1a) are replaced by plane waves (Fig. 1b) of the form

$$\mathbf{U} = \mathbf{U}_0 e^{i(k_z z \pm k_y y)}. \tag{3}$$

A significant parameter of the problem is the angle α between the wave vector $\mathbf{k}$ and the z axis. Evidently, we

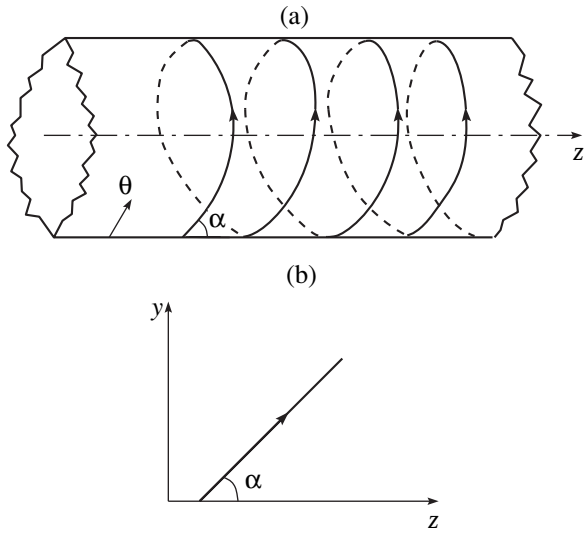


Fig. 1. (a) Helical wave for $-\infty < \theta < \infty$ and (b) the equivalent plane wave in a plate.

have $k^2 = k_y^2 + k_z^2$, $k_z = k\cos\alpha$, and $k_y = k\sin\alpha$. Then, for a helical wave, we obtain

$$\sin\alpha = \frac{\nu}{ka}. \tag{4}$$

Note that, for periodic solutions to Eq. (1), there exists a discrete set of angles α_n satisfying the relation

$$\sin\alpha_n = \frac{n}{ka}. \tag{5}$$

From Eqs. (4) and (5), it follows that, when $\nu > ka$, the angle α is an imaginary quantity and a helical wave is inhomogeneous in the angle θ; when $n > ka$, real periodic solutions are absent.

The quantity **k** is the wave number of free waves of the plate. For example, for a homogeneous elastic medium, the free waves include one longitudinal and two transverse waves. For a homogeneous elastic plate, four free waves are possible: one longitudinal wave, one shear wave, one homogeneous flexural wave, and one inhomogeneous flexural wave.

To obtain a dispersion equation with respect to the quantity k, we substitute solution (3) into equation (1) and then introduce the quantities k and α in this equation by using the aforementioned substitutions. As a result, we obtain a matrix equation

$$A(k, \alpha)\mathbf{U}_0 = 0, \tag{6}$$

where $A(k, \alpha)$ is a 3×3 matrix with the components

$$A_{11} = (1 - c\sin^2\alpha)X^2 - 1;$$

$$A_{12} = A_{21} = \frac{c}{2}\sin 2\alpha X^2;$$

$$A_{13} = -A_{31} = i\sigma\cos\alpha\frac{X}{Y};$$

$$A_{22} = (1 - c\cos^2\alpha)X^2 - 1;$$

$$A_{23} = -A_{32} = i\sin\alpha\frac{X}{Y}; \quad A_{33} = \frac{(HY)^2}{12}X^4 + \frac{1}{Y^2} - 1.$$

Here, to represent Eq. (6) in dimensionless form, the following notations were introduced:

$$X = \frac{k}{k_p}; \quad Y = k_p a; \quad H = \frac{h}{a}; \quad c = \frac{1+\sigma}{2}.$$

The dispersion equation can be represented in the form

$$\Delta = \det A(k, \alpha) = 0. \tag{7}$$

Before proceeding to the results of solving equation (7), it is necessary to make some comments.

This equation determines the values of the wave numbers of free waves in an infinite plate possessing a number of specific properties. The quantity α is an arbitrary parameter characterizing the anisotropy of the plate properties because of the bending of the shell in one plane. All three displacements of the plate under consideration are related to each other, while for an ordinary plate, transverse displacements (w) characterizing its flexural vibrations do not depend on the shear-longitudinal displacements (u, v) that occur in its plane. Solutions to Eq. (7) should strongly depend on the parameter Y; when $Y \longrightarrow \infty$, the plate under study transforms to an ordinary isotropic plate (k is independent of the angle α) and the displacement w becomes independent of u and v. (For this case, the free waves were mentioned above.)

Solutions to Eq. (7) were obtained for the following parameters: $\sigma = 0.28$, $H_1 = 0.1$, and $H_2 = 0.05$. The dimensionless wave numbers X of the free waves of the shell were determined as functions of the frequency parameter Y for several discrete values of angle α. The results of calculations are presented in Figs. 2–5.

Figure 2a shows the frequency characteristics for $\alpha = 0$ (the waves propagate along the z axis). This is the well-known case of axially symmetric waves with $n = 0$.

One can see that, in the limit of $Y \longrightarrow 0$, two waves are present with real values of wave numbers: a shear wave with $X_1 = \frac{k_t}{k_p} = \sqrt{\frac{2}{1-\sigma}} = 1.667$ and a so-called rod-type (Young's) longitudinal wave with $X_2 = \frac{k_Y}{k_p} = \sqrt{\frac{1}{1-\sigma^2}} = 1.042$. Since these two waves exist independently of one another, their branches may intersect. Since Eq. (7) is an eighth-order algebraic equation in X, the same range should contain four other roots which

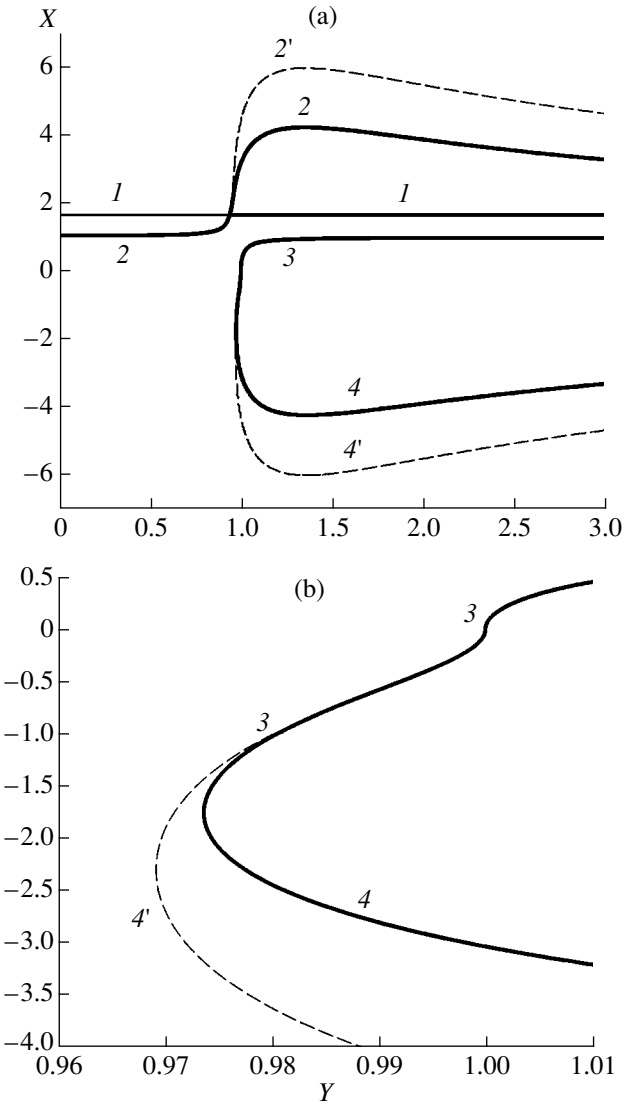


Fig. 2. (a) Dispersion curves for a cylindrical shell with the parameters $\sigma = 0.28$ and $H = 0.05$ (dashed line) and 0.1 (solid line); $\alpha = 0°$; the values $X < 0$ correspond to imaginary branches. (b) The same as in Fig. 2a for imaginary branches only (the notation is the same).

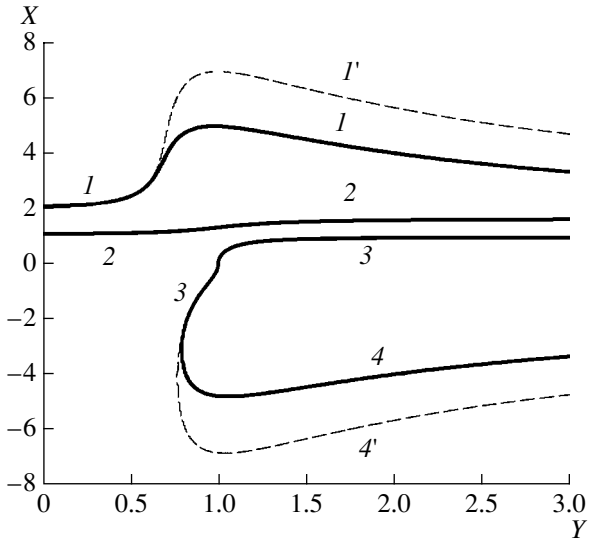


Fig. 3. Same as in Fig. 2a for $\alpha = 30°$.

belong to flexural waves and have the form $X = \pm x \pm y$, where x and y are real numbers (they are insignificant for our consideration).

At $Y = 1$ (the case of the so-called ring resonance), a longitudinal shell wave X_3 is formed; when $Y \longrightarrow \infty$, this wave becomes transformed to a longitudinal plate wave. Simultaneously, two flexural waves appear: wave *2* with a real value of X_2 (into which the rod-type longitudinal wave is transformed) and wave *4* with an imaginary wave number X_4. (In all figures, the imaginary values of wave numbers correspond to the lower half-plane in X.) The dashed lines represent the wave numbers X_2' and X_4' corresponding to $H = H_2$. One can see that branches X_2 and X_2', as well as X_4 and X_4', bifurcate at $Y > 1$, which testifies to their flexural nature. Branches X_1 and X_2 in the region $Y < 1$, as well as branches X_1 and X_3 in the region $Y > 1$, coincide for both values of H; i.e., they are practically independent of this parameter.

Figure 2b shows branches *3*, *4*, and *4'* on an enlarged scale. One can see that the latter two branches merge at $Y_0 \cong 0.98$, which suggests that the imaginary branch in the region $Y_0 < Y < 1$ should be ascribed to the longitudinal wave type.

Thus, for $Y > 1$, we have four free waves (backward waves are not considered), which corresponds to the degree of equation (7).

Subsequent figures illustrate mutual transformations of wave numbers of free waves as the angle α varies.

Figure 3 presents the corresponding data for $\alpha = 30°$. One can see a transformation of branches from one type of wave to another, as compared to the case $\alpha = 0°$. At $Y \approx 1$, the Young's branch X_2 is transformed to the shear branch X_1, and the latter is transformed to the real flexural branch X_1. Such a transformation manifests itself at all angles $\alpha > 0$. In addition, the values of the wave numbers (phase velocities) of free waves also change. A further increase in the angle α ($\alpha = 60°$ in Fig. 4) does not lead to any qualitative changes.

Figure 5 represents the limiting case $\alpha = 90°$ (the wave propagates in a circle). Here, the quantity $X = \dfrac{k}{k_p}$ is replaced by $\hat{X} = \dfrac{k}{k_u}$, where k_u is the wave number of flexural waves in a flat plate. An evident relation exists

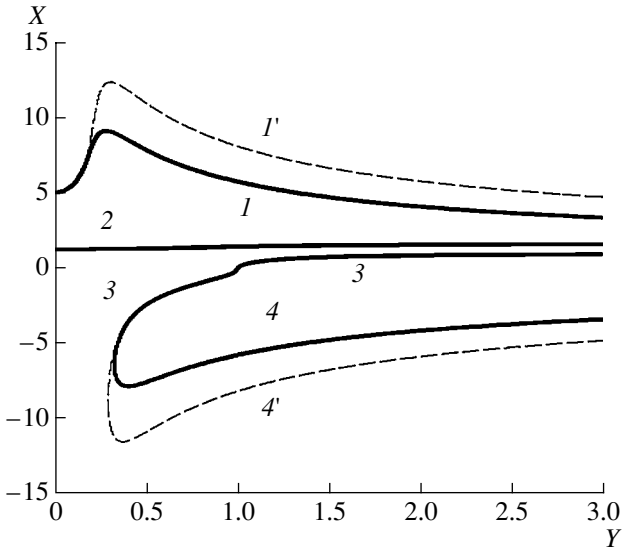


Fig. 4. Same as in Fig. 2a for $\alpha = 60°$.

between these quantities:

$$\hat{X} = X\sqrt{\frac{YH}{\sqrt{12}}}.$$

From Fig. 5, one can see that, at $Y \cong 0$, we have $\hat{X}_2 \approx 2.8$; i.e., in this frequency band, the velocity of a flexural wave propagating in a bent plate is smaller than the corresponding velocity in a flat plate. In addition to this flexural wave, there exist one purely shear wave $\hat{X}_1$ and two imaginary waves $\hat{X}_3$ and $\hat{X}_4$ (the same as in Figs. 3 and 4), which, however, appear at a frequency of $Y = 0.075$. In this case, already at $Y \approx 0.5$, the real and imaginary branches of flexural waves differ little from flexural waves of a flat plate.

An important characteristic of helical waves is the degree of anisotropy of the wave properties of the equivalent plate. Let us represent the solutions to Eq. (7) in another form. In Fig. 6, the abscissa axis represents the quantity $X_z = \frac{k_z}{k_p}$ and the ordinate axis, the quantity $X_y = \frac{k_y}{k_p}$ for individual values of the dimensionless frequency Y indicated in the figures. The curves corresponding to different dispersion branches are marked with the same numbers as in Figs. 2–5 (except for the real flexural branches marked as *1a*). This kind of representation of the results yields additional information on the properties of waves: the length of the radius-vector extending from the origin of coordinates to some of the curves corresponds to the value of X, and the angle between the radius-vector and the X_z axis is equal to α. The figures reveal the behavior of X as a function of continuously varying α.

For isotropic media, the wave numbers of free waves do not depend on the angle of their propagation (with respect to some axis), and, hence, for such media, the curves analogous to those presented here are parts of a circle. In our case, these curves may considerably

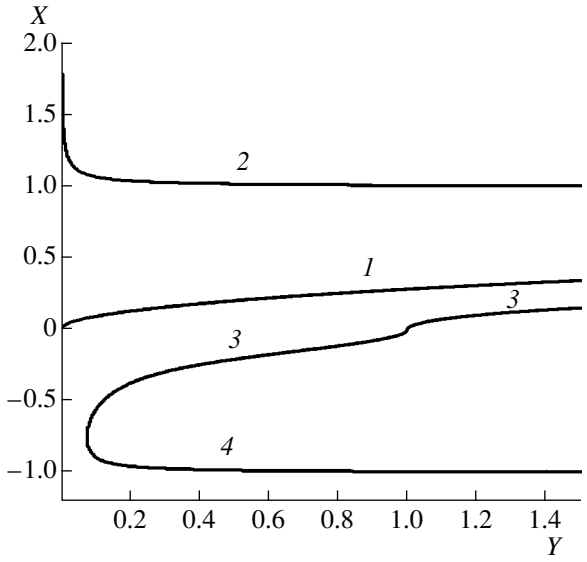


Fig. 5. Same as in Fig. 2a for $\alpha = 90°$; $\hat{X} = X\frac{k_p}{k_u}$.

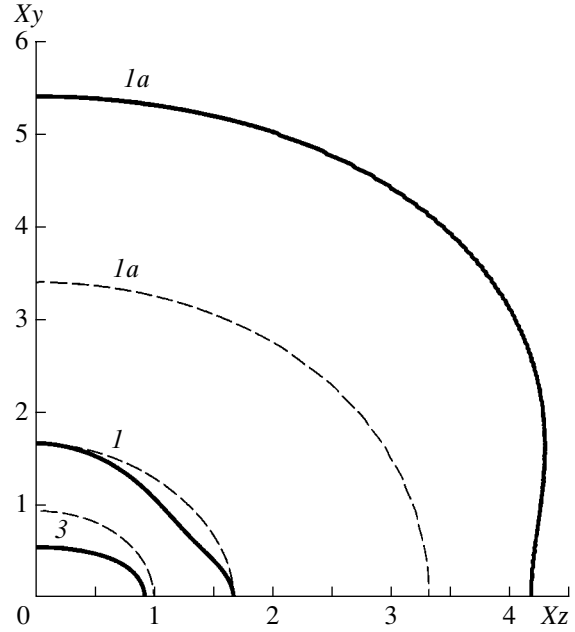


Fig. 6. Example of anisotropic properties of the shell for $H = 0.1$ and $Y = 1.2$ (solid lines) and 3 (dashed lines).

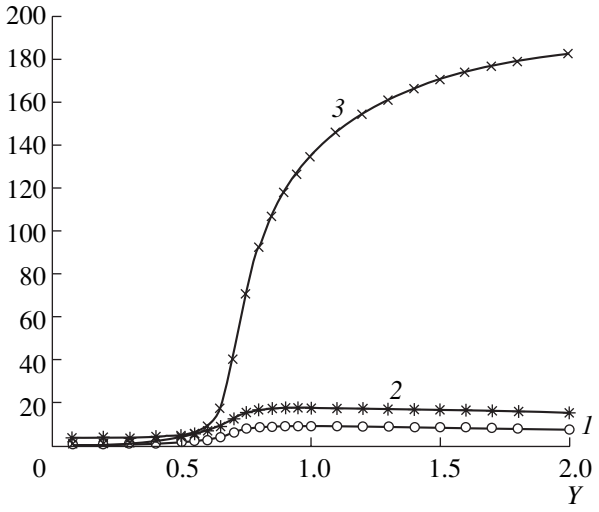


Fig. 7. Displacement amplitudes in a helical wave for the branch $X = X_1$ at $\alpha = 30°$ and $H = 0.05$.

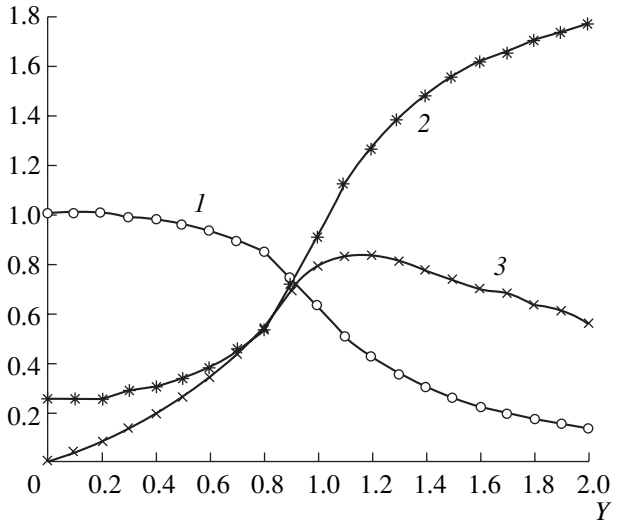


Fig. 8. Same as in Fig. 7 for the branch $X = X_2$.

deviate from a circle, and the greater this deviation is, the stronger the anisotropic properties of the shell manifest themselves. As the parameter Y grows, these properties disappear and the curves become circular.

It is of interest to consider the displacements in helical waves. For this purpose, we express the longitudinal displacements U_n, which are perpendicular to the wave front; the transverse displacements U_t, which are parallel to it; and the displacements W normal to the plate plane through the displacements u_0, v_0, and w_0 introduced earlier, the latter being related to each other by Eq. (6). Then, we easily obtain the expressions

$$\begin{aligned} U_n &= u_0\cos\alpha + v_0\sin\alpha; \\ U_t &= -v_0\cos\alpha + u_0\sin\alpha; \quad W = w_0. \end{aligned} \tag{8}$$

Without a loss of generality, we can set $u_0 = 1$ and determine the quantities v_0 and w_0 from the first two equations of system (6):

$$u_0 = 1; \quad v_0 = \frac{A_v}{A_0}; \quad w_0 = \frac{A_w}{A_0}, \tag{9}$$

where

$$A_v = -(A_{11}A_{23} - A_{12}A_{13}); \quad A_w = -(A_{11}A_{22} - A_{12}^2);$$

$$A_0 = A_{12}A_{23} - A_{22}A_{13}.$$

Substituting expressions (9) in formulas (8), we obtain the desired quantities U_n, U_t, and W as functions of frequency Y for an arbitrary angle α.

As an example, Figs. 7 and 8 show the plots of these quantities at $\alpha = 30°$ for branches X_1 and X_2 (from Fig. 3) at $H = 0.05$. The behavior of their frequency dependences confirms that these branches belong to the aforementioned type of waves.

In closing, it should be noted that the helical waves of a cylindrical shell (the free waves of an equivalent flat plate), which are considered in this paper, serve as the basis for constructing aperiodic solutions for a cylindrical shell. These solutions should describe the waves propagating along the axis of a longitudinally or transversely fixed shell (see above), as well as waveguides in the form of elastic strips cut from a cylindrical shell along or across its axis. In the latter case, the helical waves play the role of the so-called Brillouin waves.

ACKNOWLEDGMENTS

I am grateful to Yu.I. Bobrovnitskiĭ and A.I. Boĭko for the useful discussions.

REFERENCES

1. H. A. Schenck and G. W. Benthien, Akust. Zh. **41**, 828 (1995) [Acoust. Phys. **41**, 731 (1995)].
2. V. V. Muzychenko, *Sound Diffraction by Elastic Shells* (Nauka, Moscow, 1993).
3. J. F. M. Scott, J. Sound Vibr. **125**, 241 (1988).
4. M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction*, 2nd ed. (MIT Press, Cambridge, Mass., 1986).
5. L. Haumesser, D. Decultet, F. Leon, and G. Maze, J. Acoust. Soc. Am. **111**, 2034 (2002).
6. J. D. Kaplunov and M. V. Wilde, J. Acoust. Soc. Am. **111**, 2692 (2002).
7. K. Saijyou and S. Yoshikava, J. Acoust. Soc. Am. **112**, 2034 (2002).
8. S. P. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells*, 2nd ed. (McGraw-Hill, New York, 1959; Nauka, Moscow, 1966).

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 278–287. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 337–348.

BIOLOGICAL ACOUSTICS

The editorial board of *Acoustical Physics* offers a new thematic section to the readers of the journal: Biological Acoustics. The idea was stimulated by the jubilee of Nikolaĭ Andreevich Dubrovsky, Director of the Andreev Acoustics Institute and a member of the editorial board.

Dubrovsky pioneered the development of biophysical methods in acoustics, which was started on the initiative of Academician N.N. Andreev.

We hope that the articles included in the new section will give insight into the current problems of this rapidly progressing branch of acoustics. Sections of biological acoustics are also planned to appear in subsequent issues of the journal.

Differential Velocity Thresholds for an Auditory Target Moving in the Vertical Plane

M. Yu. Agaeva and A. Ya. Al'tman

Pavlov Institute of Physiology, Russian Academy of Sciences, nab. Makarova 6, St. Petersburg, 199034 Russia
e-mail: agamu@infran.ru; altman@infran.ru

Received July 22, 2003

Abstract—Differential sensitivity to the velocity of an auditory target moving in the vertical plane and the effect of the direction of motion and of the signal spectrum on the differential thresholds are studied. The motion of the auditory target was produced by a successive switching of several loudspeakers placed on an arc. It was found that an increase in the velocity in both opposite directions (from front to rear and from rear to front) leads to a monotonic increase in the average absolute differential velocity thresholds. Regression lines obtained as a linear approximation of the average absolute differential velocity thresholds do not differ significantly. For a signal whose spectrum consists of frequencies from 4 to 12.5 kHz, the absolute differential velocity thresholds are significantly lower than those for a signal occupying a frequency band from 0.25 to 4 kHz. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Localization of sound sources is one of the conditions necessary for the spatial orientation of humans and animals in a variable environment. An important feature of the localization of sound signals is their motion in space.

The perception of motion has been mostly studied in the horizontal plane and for approaching and withdrawing auditory targets [1–5].

As for the localization of a sound source moving in the vertical plane, it is studied in less detail. One of the important characteristics of perception of a moving sound source is the perception of its velocity.

It is known that localization accuracy of a sound source moving in the vertical plane depends on the spectrum of the signal. Broadband noise pulses containing high frequencies are localized more accurately than those containing low frequencies, while signals that contain no frequencies above 4 kHz cannot be definitely localized in the vertical direction [6–8]. It can therefore be expected that the localization of a sound source moving in the vertical plane relies on the same physical parameters that create the possibility to localize a fixed time-dependent sound source.

The physical origin of these features of sound source localization is that, in natural conditions, acoustic waves reflected from the body, head, and, especially, pinna change the spectrum of the signal and create spatial indications of the vertical [9]. According to [10], acoustic waves of frequencies below 2 kHz are reflected by the body, while waves with frequencies above 4 kHz are reflected by the pinna. The external ear, especially pinna, is important for localizing sound sources. Various regions of the pinna act similarly to an acoustic delay line, introducing about 0.1- to 0.3-μs-long time delays, which create considerable phase shifts at high frequencies [11, 12]. The interference of echo signals that arrive at the entrance to the ear channel with different phases increases or decreases the signal energy depending on the frequency and elevation of the sound source. At some frequencies, maxima of the energy (peaks) are observed, at others, the energy abruptly falls (dips) [13–15]. Thus, the direction-dependent transformation of the signal spectrum is a fundamental feature of the localization of sound sources in the vertical plane, which is used by the auditory system to "calculate" the signal location.

1063-7710/04/5003-0278$26.00

The goal of this work was to study the differential sensitivity to the velocity of a sound source moving in the vertical plane and the effect of the direction of motion and of the spectrum of the acoustic stimulus on the differential thresholds.

EXPERIMENTAL

Three series of experiments were performed. The first two series studied the differential thresholds as functions of the velocity of the auditory target and of its direction. In the first series, the auditory target moved from front to rear; in the second series, from rear to front. The third series of experiments studied the effect of the spectrum of the acoustic stimulus on the differential thresholds.

The first two series of experiments involved nine 23- to 35-year-old subjects (five women and four men) with normal hearing. Three of them participated in both the first and second series. Thus, the experiments involved six people.

The first two series employed a broadband (0.1 through 10 kHz) signal produced by a G2-57 low-frequency noise generator. The intensity of the acoustic signal was 60 dB SPL.

The experimental setup contained loudspeakers with similar frequency characteristics. The first series of experiments used 13 loudspeakers; the second, 15 loudspeakers. The signals were applied to the loudspeakers sequentially, thus producing a train of noise pulses. The loudspeakers were equally spaced on an arc with a radius of 1.2 m, which was installed in the vertical plane. In the course of the experiment, a subject was sitting in an armchair under the arc. The arc was coplanar with the median plane of the subject. The first and last loudspeakers were at the level of the subject's head, in front and behind it, respectively, so that the head was at the center of the semicircle described by the arc.

The apparent motion of the sound source was created by a successive switching of the loudspeakers along the arc. This acoustic stimulus was perceived by the subject as a pulsed noise source moving along the arc.

The angular velocity of the stimulus was specified by the period of the noise sequence as

$$\omega = S/(T(n-1)),$$

where ω is the angular velocity in degrees per second, S is the angular distance along the arc ($S = 180°$), T is the onset-to-onset interval in seconds, and n is the total number of loudspeakers.

The duration of each noise burst was 20 ms in the first series of experiments and 40 ms in the second one. The pulse rise and fall times were 1 ms. Eight periods of the sequence were tested: 200, 180, 160, 140, 120, 100, 80, and 60 ms. These corresponded to the following angular velocities, which were taken as the reference ones: 75, 83.5, 93.5, 107, 125, 150, 187.5, and 250 deg/s when the auditory target moved from front to rear and 64, 71, 80, 92, 107, 128.5, 161, and 214 deg/s when the auditory target moved from rear to front.

The first two series determined the differential thresholds by the method of minimal increments or method of boundaries. The subject was sequentially tested with two signals, which moved with different velocities. The velocity of the first signal was used as a reference, while the velocity of the second (test) signal was noticeably different from that of the test signal, so that the subject could easily distinguish these signals. By gradually decreasing the difference between the reference and test velocities, it was possible to determine the test signal velocity below which the subject could no longer tell the difference between the velocities. Then, varying the test velocity, the value was found beginning with which the subject detected the difference. The velocity from which the test was started was varied arbitrarily. A detailed description of the procedure used to evaluate the differential velocity thresholds is given in [16].

The third series of experiments involved five 21- to 45-year-old people (four women and one men) with normal hearing.

The signal had the form of noise pulses spanning the frequency range from 250 to 4000 Hz or from 4000 to 12 500 Hz. The signal bandwidth was adjusted by two Robotron 04024 bandpass filters for low and high frequencies with a rejection slope of 12 dB/octave. The noise signal was generated by a G2-57 low-frequency signal source and applied to two modulators. The experiments used 53 loudspeakers placed on an arc in the vertical plane. As in the previous series, the signals were applied to the loudspeakers sequentially thus forming a train of noise pulses.

Two reference velocities of 58 and 115 deg/s corresponded to the pulse repetition periods of 60 and 30 ms. The noise pulses were 40 and 15 ms long, respectively. The measurements of this series used the "staircase" or "up–down" method, a modification of the minimal increment method [17].

The acoustic stimulus changed not only its velocity of motion, but simultaneously changed its rhythmic pattern associated with the period of noise pulses. Subjects could use these changes as cues to discriminate the signals in velocity. Accordingly, an additional series of experiments was performed to analyze the differential sensitivity to the pulse repetition period. The increment of the pulse repetition period of the test signal with respect to that of the reference signal was detected. These experiments involved three subjects (nos. 1, 2, and 5).

The first loudspeaker (0° in azimuth and 0° in elevation), located on the midline of the subject's head, transmitted a signal containing 15 noise pulses. Each pulse was 40 ms long. The rise and fall times were 1 ms. The following pulse repetition periods were used as a reference: 200, 180, 160, 140, 120, 100, 80, and 60 ms,

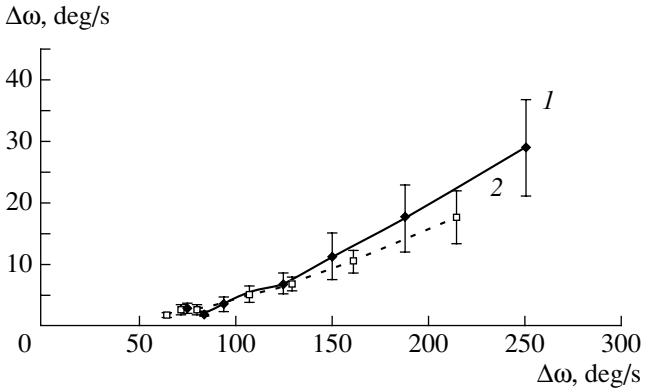


Fig. 1. Average absolute differential velocity threshold versus the reference velocity for the auditory target moving (*1*) from front to rear and (*2*) from rear to front.

which corresponded to the repetition periods used to create the apparent motion of the sound source. The differential thresholds in the repetition period were determined by the minimal increment method.

It should be noted that, in these experiments, the distance to the moving auditory target was constant and the angle subtended by the arc, along which the target moved, was 180°. Therefore, a change in the velocity of the auditory target changed the overall duration of the signal, which the subjects could use as a cue to estimate the velocity of the auditory target. Therefore, the differential sensitivity to signal duration was measured for three signal lengths: 3.1, 1.6, and 0.9 s. The signal was transmitted by the first loudspeaker, which was in front of the subject's head at a distance of 1.2 m.

The differential signal duration thresholds were measured by the up–down technique. These thresholds were compared with the differential thresholds obtained from the data on the threshold velocity discrimination. The duration of the signal was 3.1 s for the velocity of 58 deg/s, and 1.59 s for 115 deg/s.

Each experiment took about 45 min and did not visibly tire the subjects. Prior to the experiment, all subjects were trained in audition of the signals and in velocity comparison.

RESULTS

It was found that, when the auditory target moves from front to rear, the absolute differential threshold monotonically increases with velocity in the main part of the velocity range studied, as shown in Fig. 1 (curve *1*). Contiguous points of this curve were discriminated with a high degree of certainty ($p < 0.05$). An exception from the general behavior was observed in thresholds at the velocities of 75 and 83 deg/s. The threshold at 75 deg/s was higher than at 83 deg/s in all subjects. However, the differential thresholds in these two cases were not significantly distinguishable (ANOVA $F(1; 118) = 2.327$, $p > 0.05$).

The behavior of the differential threshold versus the velocity could be approximated by a linear function with a slope of 0.14. The correlation coefficient was 0.98.

When the auditory target moved from behind the subject's head forward, the average differential thresholds also increased monotonically with velocity. The differential thresholds, averaged over a group of six subjects, are shown in Fig. 1 (curve *2*). Contiguous points on this curve were discriminated reliably ($p < 0.05$). The behavior of the differential threshold versus the velocity was approximated by a linear function with a slope of 0.1. The correlation coefficient was 0.98.

It should be noted that the behavior of the differential threshold versus the velocity for the target moving from front to rear and from behind the subject's head forward showed individual distinctions. Two subjects (nos. 5 and 6) demonstrated noticeably higher differential thresholds than the others. The basic behavior observed in the group curve, however, remained the same.

Since three of the subjects (nos. 1, 2, and 5) took part in the experiments with the auditory target moving in both directions, it is reasonable to use them in the study of the effect of the direction of motion on the differential velocity sensitivity. The average absolute differential velocity thresholds for the target moving in both directions are given in Fig. 2 for each of the three subjects. The solid circles show differential thresholds for the target moving from front to rear; open squares, from rear to front. Straight lines show the linear approximation of the experimental data. The greatest correlation coefficient was obtained in the linear approximation of the absolute average differential threshold as a function of the velocity of the auditory stimulus. The correlation coefficient (R^2) was 0.964, 0.951, and 0.991 for the target moving from front to rear and 0.995, 0981, and 0.924 for the target moving from rear to front for subjects nos. 1, 2, and 5, respectively. The regression lines for the auditory stimulus moving in the two directions are not reliably distinguishable for each of the three subjects ($t_{10} = 0.152$, 0.035, and 0.114 for subjects nos. 1, 2, and 5, respectively; $p > 0.05$).

We can thus suppose that the direction of motion of an auditory target in the form of the chosen auditory stimulus exerts no effect on the differential velocity discrimination of the auditory target moving in the vertical plane.

The velocity resolution for the source moving in the vertical plane was higher in the case of the acoustic signal occupying the frequency band from 4 to 12.5 kHz.

At a velocity of 58 deg/s, the absolute differential thresholds for the signal occupying the frequency band from 0.25 to 4 kHz were higher by a factor of approximately 2 in all five subjects. At a velocity of 115 deg/s, the differential thresholds for this signal were also higher (by a factor of approximately 1.6) than for the high-frequency auditory target (with a frequency band from 4 to 12.5 kHz).

Figure 3 illustrates the dependence of the absolute differential threshold on the signal frequency band for two velocities. The difference between the thresholds for the low-frequency signal and for the signal containing high-frequency components is significant ($p < 0.01$) for the velocities of 58 and 115 deg/s. The comparison was performed by the nonparmetric Wilcoxon method.

For the signal that occupies the frequency band from 4 to 12.5 kHz, the lowest differential threshold at a velocity of 58 deg/s was 2 deg/s (subject no. 12); at a velocity of 115 deg/s, 6.5 deg/s (subject no. 13). For the signal occupying the range from 0.25 to 4 kHz, the lowest differential threshold was 4 deg/s (subject no. 12) at a velocity of 58 deg/s, and 10 deg/s (subject no. 10) at a velocity of 115 deg/s. The highest differential thresholds were observed in subject no. 11: 5.4 deg/s for the high-frequency signal and 12.7 deg/s for the low-frequency signal at a velocity of 58 deg/s and, respectively, 12.7 and 21.2 deg/s at a velocity of 115 deg/s.

According to data available from the literature, velocity thresholds for high-frequency signals are roughly the same as those for broadband signals [15]. However, in this work, the average differential velocity thresholds for a 4–12.5-kHz signal were found to be much higher than the average differential velocity thresholds obtained in the first two series of experiments with a 0.1–10-kHz auditory target moving in both directions as shown in Fig. 4. It should be noted that, although the differential velocity thresholds for a signal containing frequencies higher than 4 kHz were within the thresholds in both directions (except for subject no. 11), they were near the differential thresholds observed in subject no. 5, which were rather high. For comparison, the thresholds for subject no. 5 at a velocity of 107 deg/s were 7.3 deg/s for the stimulus moving from front to rear and 6.9 deg/s for the opposite direction. The average differential threshold for the 4 through 12.5 kHz signal for subject nos. 10, 12–14 was 7.5 deg/s at a velocity of 115 deg/s.

The distance between the loudspeakers was different: 15 deg for the target moving from rear to front and 12.8 deg for the opposite direction. In the third series of experiments, the distance between the loudspeakers was 3.6 deg, which is smaller than the minimum distinguishable angle [18, 19]. Therefore, we can assume that these conditions could affect the result of the velocity comparison. To test this assumption, we calculated the differential thresholds for the repetition period of 60 ms, which was used in all three series of experiments, so that the repetition period of acoustic pulses

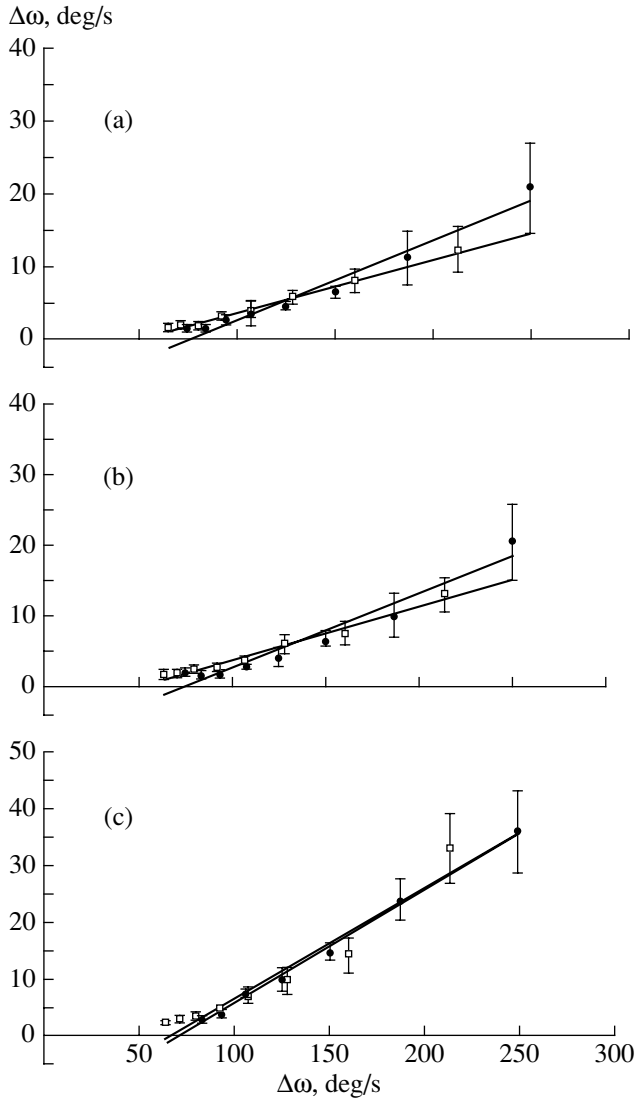


Fig. 2. Average absolute differential velocity threshold versus the reference velocity for subject nos. (a) 1, (b) 2, and (c) 5 with the auditory target moving (solid circles) from front to rear and (open squares) from rear to front. Straight lines represent the result of the linear approximation.

was the same, while the velocity was varied by varying the angular separation between the loudspeakers. It was found that subjects failed to reliably discriminate the differential thresholds in a repetition period when the angular separation between the loudspeakers was 12.8 and 15 deg ($p > 0.5$, $F(1; 10) = 1.53$). The thresholds were reliably discriminated when the angular separation was 3.6 and 12.8 deg ($p < 0.5$, $F(1; 9) = 6.11$) and 3.6 and 15 deg ($p < 0.5$, $F(1; 8) = 7.06$). Thus, the angular separation between the loudspeakers affects the differential thresholds if it is greater than 3.6 deg.

Since the velocity of the auditory target was varied by varying the repetition period, when comparing the velocities, the subjects could utilize not only the change

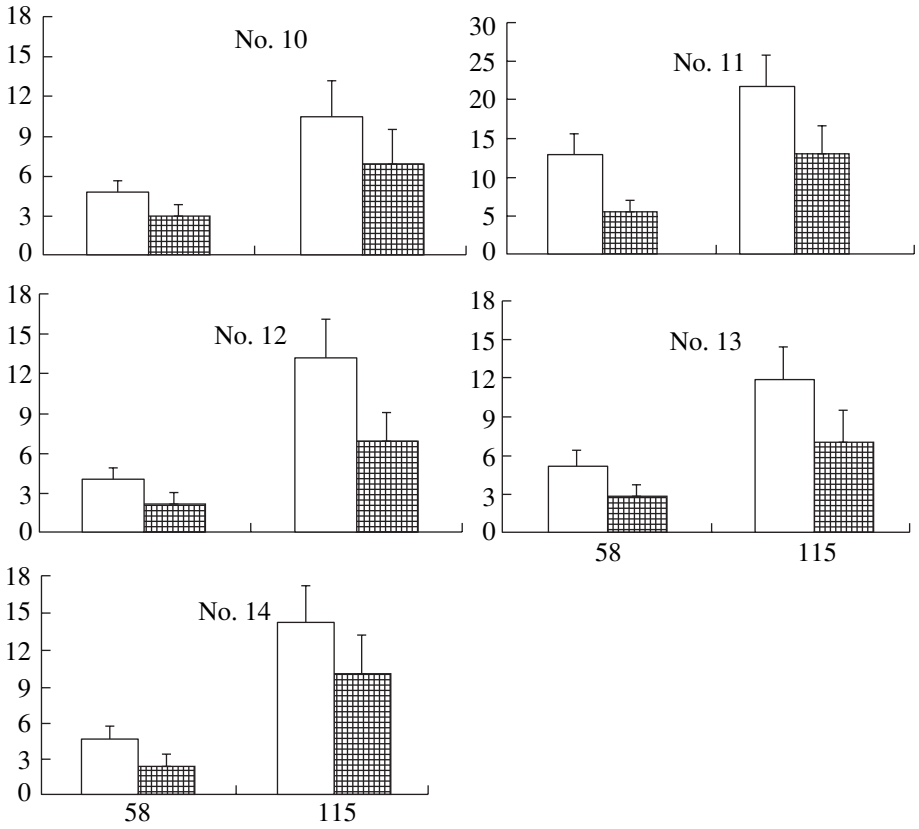


Fig. 3. Absolute differential velocity threshold versus the reference velocity for subject nos. 10–14 for the acoustic signals occupying frequency bands from (open rectangles) 0.25 to 4 kHz and (hatched rectangles) from 4 to 12.5 kHz.

in the velocity itself but also the change in the signal rhythm. To test this hypothesis, we converted the differential velocity thresholds into the corresponding thresholds in repetition period (ΔT) and compared them with differential thresholds in the repetition period for a stationary signal. The relative differential thresholds $\Delta T/T$ in velocity and in the rhythm of the signal for subjects nos. 1, 2, and 5 are shown in Fig. 5.

Subject no. 1 did not show a reliable difference ($p > 0.05$) in the relative thresholds in repetition period when the auditory target moved from front to rear at T = 200, 140, 100, and 80 ms and from rear to front at T = 140, 120, 100, and 60 ms.

For subject no. 2, the relative differential thresholds were not reliably distinguishable ($p > 0.05$) for the target moving from front to rear at repetition periods of T = 180 and 160 ms. For the target moving in the opposite direction, the differences were reliable ($p < 0.05$) for all repetition periods.

For subject no. 5, the relative differential thresholds were not reliably distinguishable ($p > 0.05$) for the motion from front to rear at repetition periods of T = 180 and 160 ms and, for the motion in the opposite direction, at a repetition period of 180 ms. In other cases, the relative differential velocity thresholds in both directions differed from the relative differential thresholds in repetition period and were higher.

In our experiments, the trajectory of the auditory target was fixed: it had the form of an arc enclosing an angle of 180 deg. Therefore, when the velocity changed, the only parameter that changed was the total time necessary for the sound stimulus to travel this distance. The differential sensitivity was measured with sound durations of 3, 1.6, and 0.9 s. The differential thresholds were reliably ($p < 0.05$) distinguished for each duration. The relative differential thresholds in duration, averaged over a group of five subjects, are shown in Fig. 6 (curve *1*). For the sake of comparison, curves *2* and *3* in this figure represent the differential sensitivity in sound duration. Curve *2* is borrowed from [1]; curve *3*, from [20]. As can be seen from Fig. 6, the relative duration thresholds obtained in this work actually coincide with those reported in the papers cited above.

The relative differential thresholds in sound duration were compared with converted relative differential thresholds in total sound duration for a moving auditory target. The relative duration thresholds for 3.1 and 1.6 s were compared with converted relative thresholds for velocities of 58 and 115 deg/s obtained with the low-frequency and high-frequency signals, which corre-

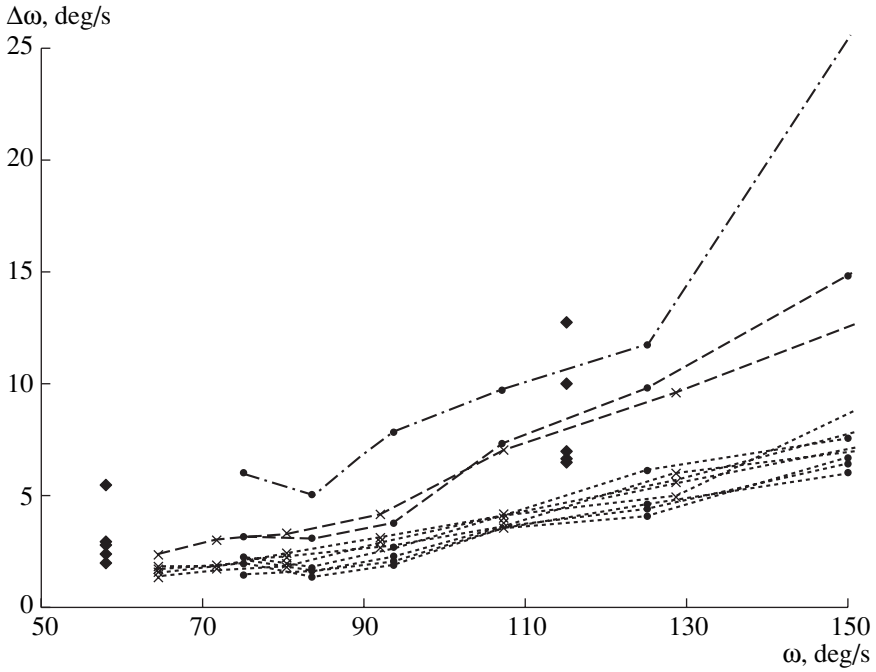


Fig. 4. Differential velocity threshold versus the reference velocity (solid diamonds) for a high-frequency (4–12.5 kHz) signal with the target moving (solid circles) from front to rear and (stars) from rear to front. The dashed lines represent the differential velocity thresholds for the auditory target moving in opposite directions: (long dashes) subject no. 5, (dot-and-dash line) subject no. 6, and (short dashes) the rest of the subjects.

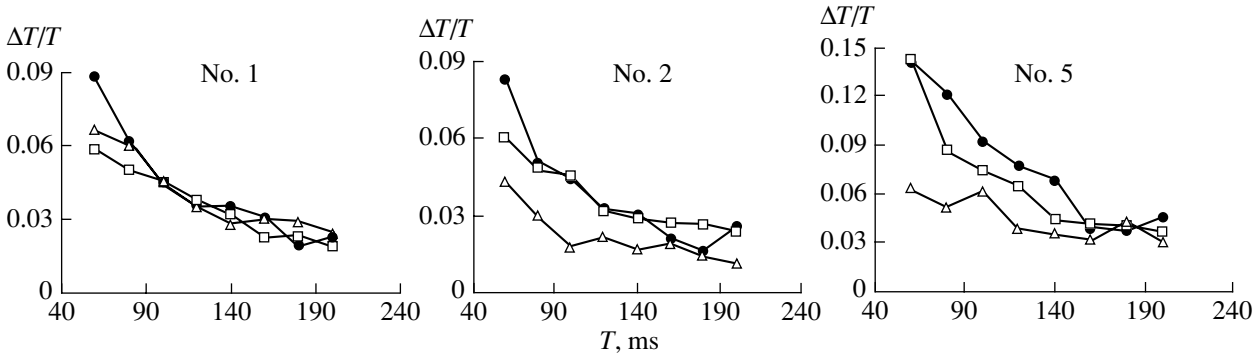


Fig. 5. Relative differential threshold in the repetition period for fixed and moving signals versus the repetition period for subject nos. 1–5: (full circles) the auditory target moves from front to rear and (open squares) in the opposite direction. The open triangles refer to the discrimination of the repetition period.

sponded to durations of 3.1 and 1.6 s. The difference between differential thresholds for moving and fixed sources was also significant for the signal spread over a frequency band of 0.25–4 kHz, as well as 4–12.5 kHz ($F(3, 158) \geqslant 2.66$, $p < 0.05$). Differential thresholds for signals with a frequency band from 0.25 to 4 kHz and from 4 to 12.5 kHz are indicated in Fig. 6 by numbers *4* and *5*, respectively.

Thus, differential thresholds in sound duration for a fixed source are significantly different from the converted duration thresholds for a moving source and are substantially higher. From the above results, one can conclude that the differential sensitivity to velocity of the auditory target in the vertical plane did not depend on variations in the total sound duration.

DISCUSSION

Below, we discuss the following issues:

(1) The difference between the recognition of a moving rhythmic signal and recognition of the repetition rate of acoustic signals, because our study used apparent motion, which was produced by changing the repetition period (and, therefore, the rate) of noise pulses. Also, the angle subtended by the trajectory was constant (180 deg) and the total duration of the motion varied; therefore, the results were compared with the

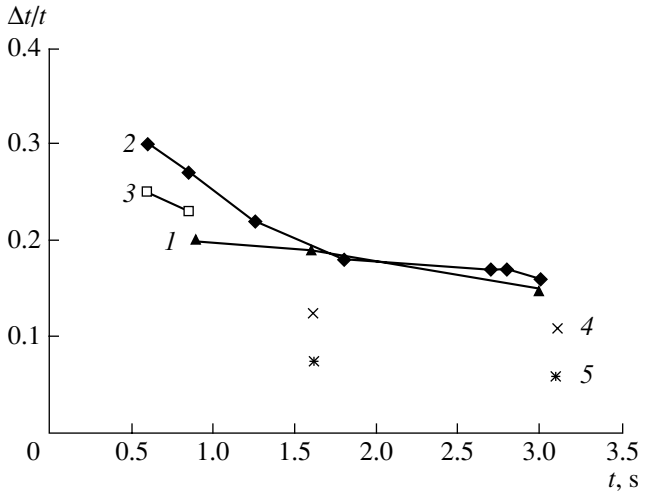


Fig. 6. Relative differential threshold in signal duration versus the duration: results (*1*) obtained in this work and (*2*, *3*) borrowed from [1, 20] (respectively). The crosses and asterisks show the converted relative thresholds for moving auditory targets with frequency bands from 0.25 to 4 kHz and from 4 to 12.5 kHz, respectively.

differential sensitivity to duration of the acoustic stimulus.

(2) The effect of the direction of motion on the differential velocity sensitivity in the vertical plane. The auditory target was moved in the experiments from front to rear and from rear to the front.

(3) The effect of the spectrum of sound on the differential sensitivity to the velocity of the auditory target.

(1) Our experiments used apparent motion of the auditory target. When the motion was formed, the velocity and repetition period of the signal produced by sequentially switching the loudspeakers were directly related. The velocity of the motion changed with an increase or decrease in the repetition period or rate of the series of noise signals, and the subjects could evaluate the change in the velocity from variations in the signal periodicity. The subjects could also rely in their estimates on the variation in the temporal gap between the stimuli, because the pulse length was constant.

We first consider the possibility of discriminating the velocity from the pulse repetition rate. According to our results, it can be assumed that subject no. 1 estimated a change in the auditory target velocity mostly from a change in the repetition period. The relative differential thresholds in the problem of discriminating the repetition period reliably coincided with converted increments of the repetition period in the problem of discriminating the velocity of the auditory target at a period of 200 and 140–60 ms for a target moving from front to rear and at a period of 140–100 and 60 ms for a target moving in the opposite direction. Subject nos. 2 and 5 compared the velocities based on the change in the rhythm of the signal only when the repetition period was longer than 140 ms. At shorter repetition periods, the relative differential thresholds in the discrimination of the repetition period were significantly different and were lower than those in the velocity discrimination by a factor of approximately 1.7.

The information on boundary conditions in forming an impression of motion can be found in the literature. The boundary conditions for creating a continuous moving auditory target in the horizontal plane under dichotic stimulation were studied in [21]. The threshold click repetition rate for the motion to be perceived was 7.6/s. In our experiment, the repetition rate was 7.14/s at a repetition period of 140 ms, and 8.3/s at a period of 120 ms.

It was found that the arrangement of loudspeakers in the horizontal or vertical plane and their number (two or three) exerts almost no effect on the impression of continuous apparent motion [22]. The greatest effect is produced by durations of the sound stimulus and the interval between the stimuli. The critical interval between periodic acoustic pulses that is necessary to create an impression of motion is between 0.1 and 0.15 s for the horizontal plane and for the source moving to or from the listener [23].

Based on the above results, we can assume that listeners can discriminate the velocity of a periodic signal on the basis of the change in the repetition period (subject no. 1) and also the change in the velocity itself (subject nos. 2 and 5). However, no matter how the subjects discriminated the velocity, there were no noticeable difference in the differential thresholds.

Let us compare our results with the information available from the literature about the discrimination of the rate at which the noise is interrupted. The results of the comparison are shown in Fig. 7a, which represents the relative thresholds in the discrimination of velocity and repetition period for subject nos. 5 and 2 in terms of the signal repetition period. The relative velocity and repetition period discrimination thresholds are shown by full circles and asterisks, respectively. Curves *1–3* show the differential sensitivity to the signal repetition rate [24–26]. To compare these data with our results, the repetition rate thresholds were converted to the corresponding repetition period thresholds. The relative repetition period discrimination thresholds and those in the noise interruption rate are comparable in their magnitude and behavior. The repetition period discrimination curve obtained for subject no. 2 is close to the data reported in [25]. The curve obtained for subject no. 5 continues passing into the curve borrowed from [26]. The velocity discrimination curves differ from the rate discrimination curves in their faster growth beginning with a period of 140 ms.

As stated above, the subjects could also rely in their velocity discrimination estimates on the change in the length of the temporal gap between sequential pulses, because the pulse repetition period was varied, while the length of the noise pulse was constant. According to the literature, the differential threshold discrimination

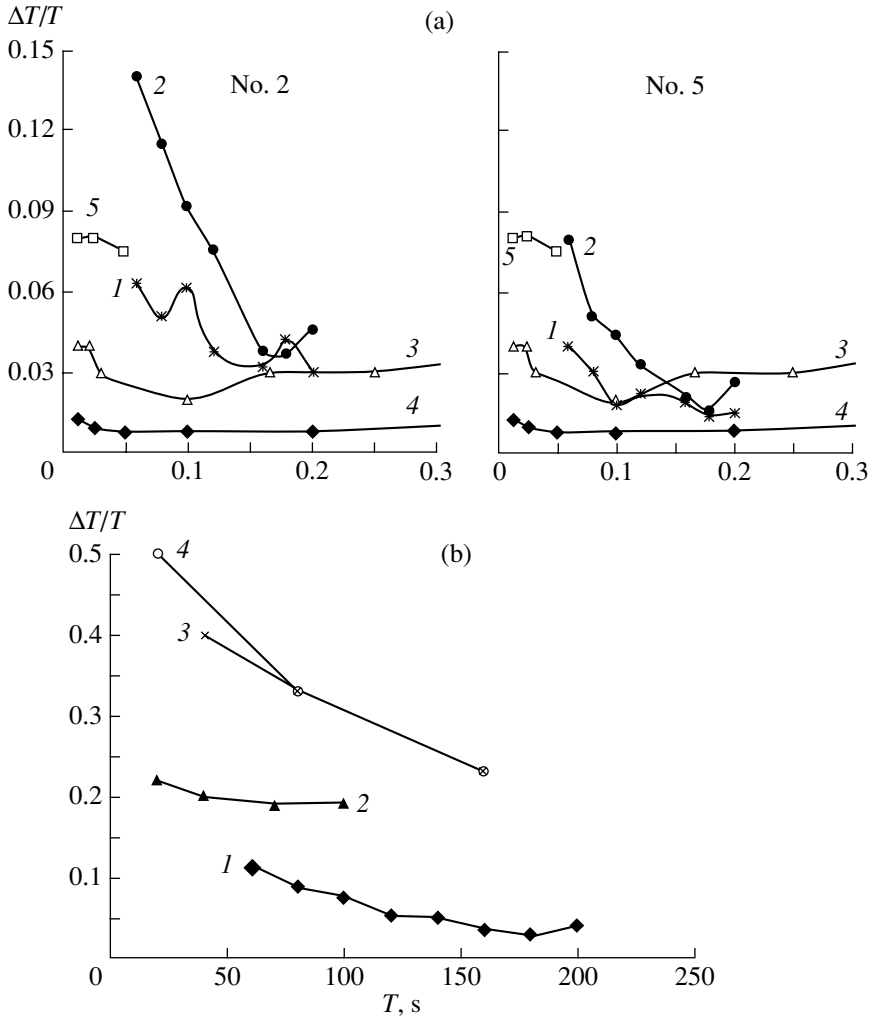


Fig. 7. (a) Relative differential repetition period threshold in the case of discriminating the velocity and repetition period of a fixed sound source for subject nos. 2 and 5 versus the repetition period: (*1*, *2*) the repetition period thresholds for a fixed source and a moving source, respectively; the thresholds in noise interruption rate borrowed from (*3*) [25], (*4*) [24], and (*5*) [26] and represented as relative differential thresholds in repetition period. (b) Relative differential thresholds in duration versus the interval between stimuli: (*1*) experimental data of the present study and data borrowed from (*2*) [29], (*3*) [28], and (*4*) [27].

of the temporal gap was roughly 40 to 20% for durations of 40 to 160 ms, respectively [27–29]. Figure 7b shows the relative thresholds in duration and repetition period for a moving auditory target. As can be seen from the figure, the increment in the repetition period when discriminating the velocity was found to be much smaller than the relative threshold in the duration discrimination. These quantities differed by a factor of 2 to 6. Thus, a change in the temporal gap between the stimuli that corresponds to the change in velocity smaller than this is necessary for its detection.

As we noted above, a change in the pulse repetition period, which determines the velocity of the auditory target, also affects the total duration of the auditory stimulus. Therefore, the subjects can be expected to discriminate the signals based on not only the velocity but also the duration of the stimulus.

To test this hypothesis, we measured the differential duration thresholds. The relative velocity thresholds were noticeably different from the duration thresholds, as shown in Fig. 6. We can therefore conclude that, with this type of signal, the subjects did not use such information as the total signal duration in their velocity estimates. This is validated by the data available from the literature. In paper [1], a train of clicks was used as a signal, and the velocity of the continuous auditory target was changed by changing the duration. The study showed that it was the velocity rather than the total duration that the subjects discriminated.

(2) The effect of the direction of motion on the differential thresholds. It could be supposed that there is a preferable direction for a listener in which the person exhibits the highest differential sensitivity. However, as we found out in this study, the differential sensitivity to the velocity of the auditory target was the same in oppo-

site directions. The regression lines, which were calculated by averaging the data obtained from six subjects for the motion from front to rear and for six subjects for the motion from rear to front, reliably coincide (t = 0.489 and $p > 0.05$). The regression lines, which were calculated by approximating the average differential velocity thresholds for the three subjects that participated in two experiments (with the auditory target moving from front to rear and in the opposite direction), also coincide with confidence. This is corroborated by results obtained for the motion in the horizontal plane and for the auditory target moving to and from the listener. For the horizontal motion, there were no reliable differences found between differential thresholds for three velocities with the target moving in opposite directions: from left to right and from right to left [30].

(3) In this study, we have shown that the differential velocity sensitivity depends on the spectrum of the sound stimulus. The differential velocity thresholds for signals that occupy the frequency band from 4 to 12.5 kHz were lower than those for signals with a frequency band from 0.25 to 4 kHz by a factor of approximately 1.5. These results are in good agreement with those obtained in localization of a fixed sound source in the vertical plane. Data obtained in [6, 7] show that signals containing high frequencies are localized with a higher accuracy than signals consisting of low frequencies (below 4 kHz). Error in localizing the signals consisting of frequencies below 3 kHz was found to be higher by a factor of 3 than that for noise pulses containing frequencies above 3 kHz [31].

Only two papers have studied the motion of a sound source in the vertical plane, particularly, the minimum audible movement angle. In [32], the signal had the form of broadband 1.8-ms-long pulses, and in the other study [33], noise was used in the frequency band from 500 to 8000 Hz. It is therefore impossible to compare the effect of the signal spectrum on the angular resolution of a sound source in the vertical plane, i.e., on the minimum audible movement angle.

It should be noted that, in this paper, although the differential velocity thresholds for a signal containing frequencies above 4 kHz were within the differential velocity thresholds for motion in opposite directions (except for subject no. 11), they were near the differential thresholds observed in subject no. 5, which were rather high. For the sake of comparison, for subject no. 5, the velocity thresholds at a velocity of 107 deg/s were 7.3 deg/s for a stimulus moving from front to rear and 6.9 deg/s for the opposite direction. The average differential threshold for a signal with a frequency band from 4 to 12.5 kHz observed in subject nos. 10 and 12–14 was 7.5 deg/s at a velocity of 115 deg/s. According to data available from the literature, the localization accuracy in the vertical plane for signals containing high frequencies is approximately the same as that for broadband signals [34]. We cannot disregard the fact that this individual sensitivity can be explained by the random selection of the group of subjects.

However, as we detailed in the "Results" section, the effect of distance between the loudspeakers on the threshold cannot be ignored. According to our results, if the angular separation of the loudspeakers is greater than 3.6 deg, it affects the differential threshold.

We intend to use the data obtained in this work to study the effect of precedence of a moving signal.

ACKNOWLEDGMENTS

We respectfully devote this paper to the jubilee of the prominent acousticial N.A. Dubrovsky, with whom we had a long-term fruitful cooperation.

This work was supported by the Russian Foundation for Basic Research, project no. 03-06-80371-a and by the "Leading Scientific Schools" project no. NSh-2230.2003.4.

REFERENCES

1. J. A. Altman and O. V. Viskov, J. Acoust. Soc. Am. **61**, 816 (1977).
2. D. R. Perrott, V. Buck, W. Waugh, and T. Z. Strybel, J. Aud. Res. **19** (4), 277 (1979).
3. V. P. Romanov, Fiziol. Cheloveka **67**, 844 (1981).
4. D. W. Grantham, J. Acoust. Soc. Am. **79**, 1939 (1986).
5. I. G. Andreeva and Ya. A. Al'tman, Sens. Sist. **15** (4), 295 (2001).
6. S. K. Roffler and R. A. Butler, J. Acoust. Soc. Am. **43**, 1255 (1968).
7. R. A. Butler and N. Planert, Percept. Psychophys. **19** (1), 103 (1976).
8. F. Asano, Y. Suzuki, and T. Sone, J. Acoust. Soc. Am. **88**, 159 (1990).
9. S. Freedman and G. Fisher, J. Aud. Res. **8** (1), 15 (1968).
10. J. Blauert, *Spatial Hearing: the Psychophysics of Human Sound Localization* (MIT Press, Cambridge, Mass., 1997; Énergiya, Moscow, 1979).
11. D. W. Batteau, Proc. R. Soc. London, Ser. B **168** (1011), 158 (1967).
12. Y. Hiranaka and H. Yamasaki, J. Acoust. Soc. Am. **73**, 291 (1983).
13. J. Hebrank and D. Wright, J. Acoust. Soc. Am. **56**, 1829 (1974).
14. J. Hebrank and D. Wright, J. Acoust. Soc. Am. **56**, 935 (1974).
15. R. A. Butler and K. Belendiuk, J. Acoust. Soc. Am. **61**, 1264 (1977).
16. M. Yu. Agaeva and N. I. Nikitin, Fiziol. Cheloveka **25** (3), 47 (1999).
17. H. Levitt, J. Acoust. Soc. Am. **730**, 467 (1970).
18. D. R. Perrott and K. Saberi, J. Acoust. Soc. Am. **87**, 1728 (1990).
19. R. B. Wettschureck, Acustica **28**, 197 (1973).
20. H. B. Ruhm and E. O. Mencke, J. Speech Hear. Res., No. 9, 371 (1966).
21. O. V. Viskov, Fiziol. Cheloveka **1** (2), 371 (1975).
22. T. Z. Strybel, A. M. Witty, and D. R. Perrott, Percept. Psychophys. **52** (2), 139 (1992).

23. Ya. A. Al'tman, *Localization of a Moving Source of Sound* (Nauka, Leningrad, 1983).
24. G. H. Mowbray, J. W. Gebhard, and C. L. Byham, J. Acoust. Soc. Am. **28**, 106 (1956).
25. I. Pollack, Am. J. Psychol. **65**, 544 (1952).
26. G. A. Miller and W. G. Taylor, J. Acoust. Soc. Am. **20**, 171 (1948).
27. L. A. Chistovich, Akust. Zh. **5**, 480 (1959) [Sov. Phys. Acoust. **5**, 493 (1959)].
28. S. M. Abel, J. Acoust. Soc. Am. **52**, 519 (1972).
29. C. D. Creelman, J. Acoust. Soc. Am. **34**, 582 (1962).
30. S. Carlile and V. Best, J. Acoust. Soc. Am. **111**, 1026 (2002).
31. R. A. Butler and R. A. Humanski, Percept. Psychophys. **51** (2), 182 (1992).
32. K. Saberi and D. R. Perrott, J. Acoust. Soc. Am. **88**, 2639 (1990).
33. T. Z. Strybel, C. L. Manligas, and D. R. Perrott, Hum. Factors **34** (3), 267 (1992).
34. M. B. Gardner, J. Acoust. Soc. Am. **54**, 1489 (1973).

Translated by A. Khzmalyan

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 288–294. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 349–356.

BIOLOGICAL ACOUSTICS

Specific Features of Vowel-Like Signals of White Whales

V. M. Bel'kovich and S. A. Kreichi

Shirshov Oceanology Institute, Russian Academy of Sciences, Nakhimovskii pr. 36, Moscow, 117851 Russia
e-mail: mmbbl@sio.rssi.ru

Received October 18, 2003

Abstract—The set of acoustic signals of White-Sea white whales comprises about 70 types of signals. Six of them occur most often and constitute 75% of the total number of signals produced by these animals. According to behavioral reactions, white whales distinguish each other by acoustic signals, which is also typical of other animal species and humans. To investigate this phenomenon, signals perceived as vowel-like sounds of speech, including sounds perceived as a "bleat," were chosen A sample of 480 signals recorded in June and July, 2000, in the White Sea within a reproductive assemblage of white whales near the Large Solovetskii Island was studied. Signals were recorded on a digital data carrier (a SONY minidisk) in the frequency range of 0.06–20 kHz. The purpose of the study was to reveal the perceptive and acoustic features specific to individual animals. The study was carried out using the methods of structural analysis of vocal speech that are employed in lingual criminalistics to identify a speaking person. It was demonstrated that this approach allows one to group the signals by coincident perceptive and acoustic parameters with assigning individual attributes to single parameters. This provided an opportunity to separate conditionally about 40 different sources of acoustic signals according to the totality of coincidences, which corresponded to the number of white whales observed visually. Thus, the application of this method proves to be very promising for the acoustic identification of white whales and other marine mammals, this possibility being very important for biology. © *2004 MAIK "Nauka/Interperiodica".*

Registration of wild animals in native habitat has many applied aspects. These are the Red Book, hunting, studying migration of various species, and many other applications. Various methods are used to monitor wild animals, starting from questioning local residents and hunters and ending with observations from ships, aircrafts, and satellites. Various animal markers are also used for this purpose, including natural specific features of shape, color, etc., and artificial ones, such as radio sets.

We systematically monitor white whales at a stationary base in the coastal zone of the White Sea employing synchronous audio and video detection [1–3, 5, 6]. An acoustic circuit consisting of a hydrophone, an amplifier, and a digital data storage medium (a SONY minidisk) provides an opportunity to record the communicative signals of white whales within the frequency range of 0.06–20 kHz.

A special feature of the observation site is the presence of a local population with a relatively stable livestock number. These are mainly adult females with calves, who permanently demonstrate playful behavior. An observation tower is erected at a small stony bank at 0.5 km from the shore and does not influence the behavior of the animals.

The stability of livestock composition was the reason for the stability of the set of observed acoustic signals, which, in the case of systematic observations, became familiar and audibly recognizable, as in the case when a human in an unfamiliar place begins gradually to recognize the voices and speech of other humans whom he meets every day. This fact suggested an opportunity to use the communicative signals of white whales as a natural acoustic marker that carries information on individual properties of the sound-emitting system of an animal.

Communicative signals of white whales are complex sound sets (Fig. 1a) consisting of different components well distinguished audibly and forming well remembered structures. Due to the stable framework of such structures, they can be combined perceptually into classes and recognized in the flow of emitted signals. An important role in recognition is played by such parameters of single structure elements as their duration, intensity, positions on the frequency and time axes, frequency bandwidth, and presence or absence of repetitions.

A communicative signal of a white whale is similar in its structure to a speech signal, which also consists of stable acoustic structures (Fig. 1b) recognizable and distinguishable by the same frequency, amplitude, and time parameters. This provides an opportunity to use the structure of a speech signal as a model for studying the communicative signals of white whales and apply the research methods used, for example, for the identification of a speaking person by the characteristics of his voice and speech. Figure 1 shows the spectral-time structure of a speech statement for comparison.

The elements forming the spectral-time structure of a communicative signal of a white whale are vocal-like (vowel-like), whistle, noise, and pulsed acoustic components containing information on the individual prop-

1063-7710/04/5003-0288$26.00

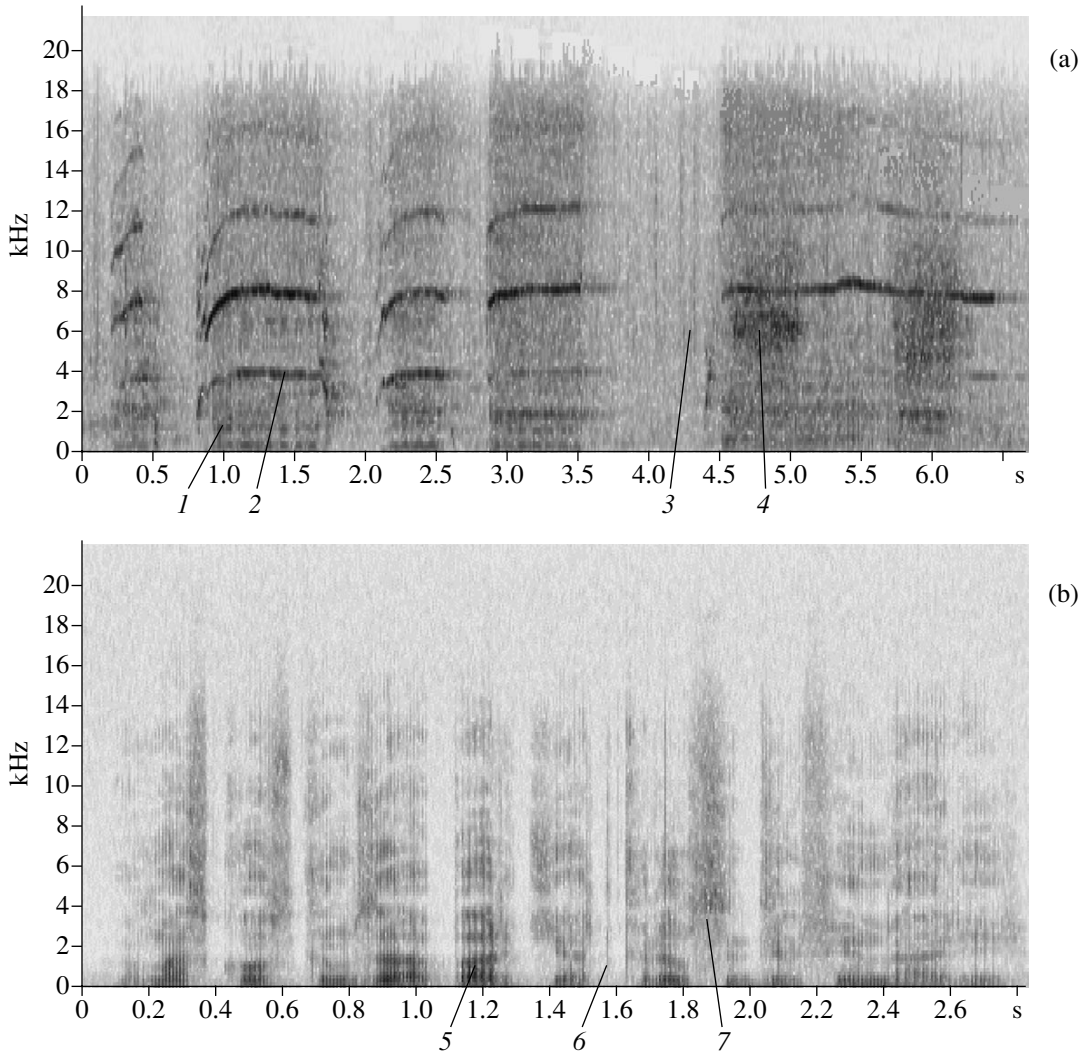


Fig. 1. (a) Spectral–time structure of a communicative signal of a white whale, which contains (*1*) vowel-like, (*2*) whistle, (*3*) pulsed, and (*4*) noise components. (b) Spectral–time structure of a speech statement: *V Moskovskom zooparke otkrylsya del'finarii* (*A dolphinarium has been opened at the Moscow Zoo*), which contains (*5*) vowels, (*6*) pulsed, and (*7*) noise components.

erties of the transmitting system of an animal, its emotional state, and the behavioral situation in which the signal under investigation was measured. As is shown in Fig. 1a, relatively simple structures can form blocks similar in composition, which, being repeated with certain changes and additions, form more complex structures analogously to the processes of human speech, when single syllables consisting, in their turn, of sound combinations form a more complex structure of a word [7]. The information contained in the elements of such structures and also the stable framework and repetition of blocks can be used (especially in the case of a simultaneous visual observation) for the formation of individual and group distinctive features, which can provide an opportunity to judge the number of heads in a drove, habitat, type of group behavior, etc.

The study of the communicative signals of white whales, which were measured during the season of

Table 1. Percentage distribution over classes of voice timbre

Male voices	High tenor	Tenor	High baritone	Baritone	High bass	Bass
Occurrence, %	5.9	22.3	26	30.8	9.5	5.5
Female voices	High soprano	Soprano	High mezzosoprano	Mezzosoprano	High contralto	Contralto
Occurrence, %	8.3	23.7	31	20	13.7	3.3

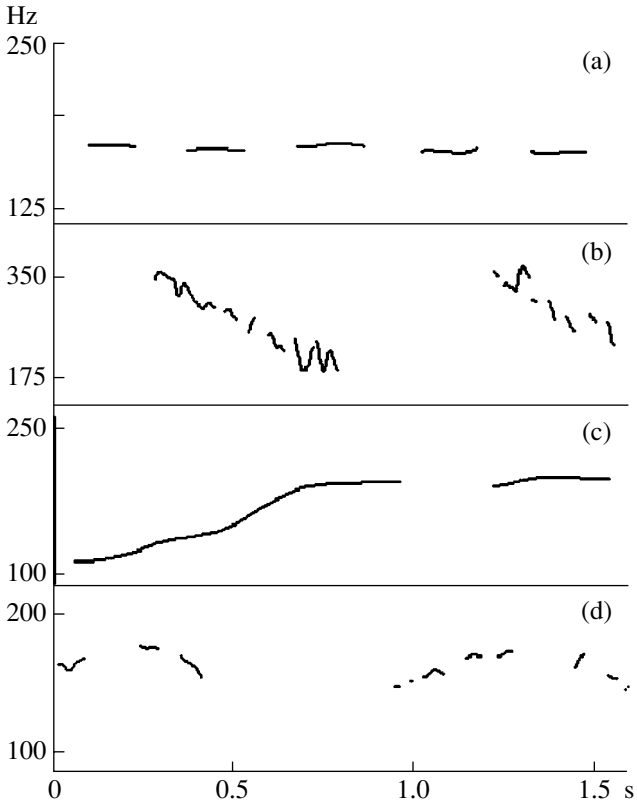


Fig. 2. Examples of shapes of frequency profiles for vowel-like signals of white whales: (a) monotonic, (b) descending, (c) ascending, and (d) arc-shaped.

observations, included perceptive and instrumental analysis. At the beginning, the perceptive characteristics of each component were determined by audio monitoring of separate signals singled out as integral spectral-time structures. For example, the tone-frequency type (clear, noisy, high, low, or medium), the type of frequency profile (monotonic; ascending; descending; or arc-shaped, in particular, convex or concave), and the presence of modulation were determined for the "vocal" component.

Instrumental acoustic analysis was conducted using the Speech Analyzer computer code, which provides an opportunity to determine amplitude–time, frequency–time, and spectral–time characteristics of the measured signals. The data of instrumental analysis confirmed the majority of perceptive observations and allowed us to evaluate quantitatively the distinctive features singled out. For example, the frequency limits were determined in the process of grouping into high, medium, and low sounds according to the high-frequency sound character of "vocalization."

As we have already indicated above, using human speech and singing as a model for investigating the communicative signals of white whales, we applied the corresponding terms. Such categories as "vocalization," "vowel-like sounds," etc. characterize only an audible result and not the mechanism of sound generation.

One of the characteristic features noted while listening to the communicative signals of white whales is the repetitive character of the blocks of a sound complex. For example, two-time repetition occurred in 20% of all signals; three-times, in 5%; four-times, in 3%; and five-times, in 1%. The characteristic feature in this case is the inverse dependence of block duration on the number of repetitions.

The types of shapes of the frequency profiles that were determined perceptively as monotonic, descending, ascending, and arc-shaped ones were confirmed graphically in the process of instrumental analysis (Fig. 2).

The occurrence of the indicated shapes of the frequency profiles of vowel-like components is distributed as follows: monotonic profiles make up 40% of the total amount; descending profiles, 50%; ascending profiles, 7%; and bow-shaped profiles, 3%.

Perceptively, two types of durations of vowel-like sounds were distinguished in communicative signals of white whales: shorter sounds perceived as vowels of human speech and longer sounds similar to vowels in singing. Figure 3 for comparison shows oscillograms of singing a Russian vowel "a" and a long vowel-like sound of a white whale with amplitude modulation. The modulation frequency of a white whale has the same order of magnitude as the vibration frequency in singing, while the modulation depth is approximately two times greater, which is perceived audibly as a "bleat." The ratio of the levels of amplitude and frequency modulation can be also considered as an individual feature of an animal.

Measurements of the modulation period demonstrated that it is within the range of 80–120 ms, which

Table 2. Average ranges of fundamental frequencies for separated timbre groups

Male voices	High tenor	Tenor	High baritone	Baritone	High bass	Bass
Fundamental frequency, Hz	260–115	230–115	190–100	160–80	140–80	120–80
Female voices	High soprano	Soprano	High mezzosoprano	Mezzosoprano	High contralto	Contralto
Fundamental frequency, Hz	400–220	360–220	320–170	270–150	250–130	240–110

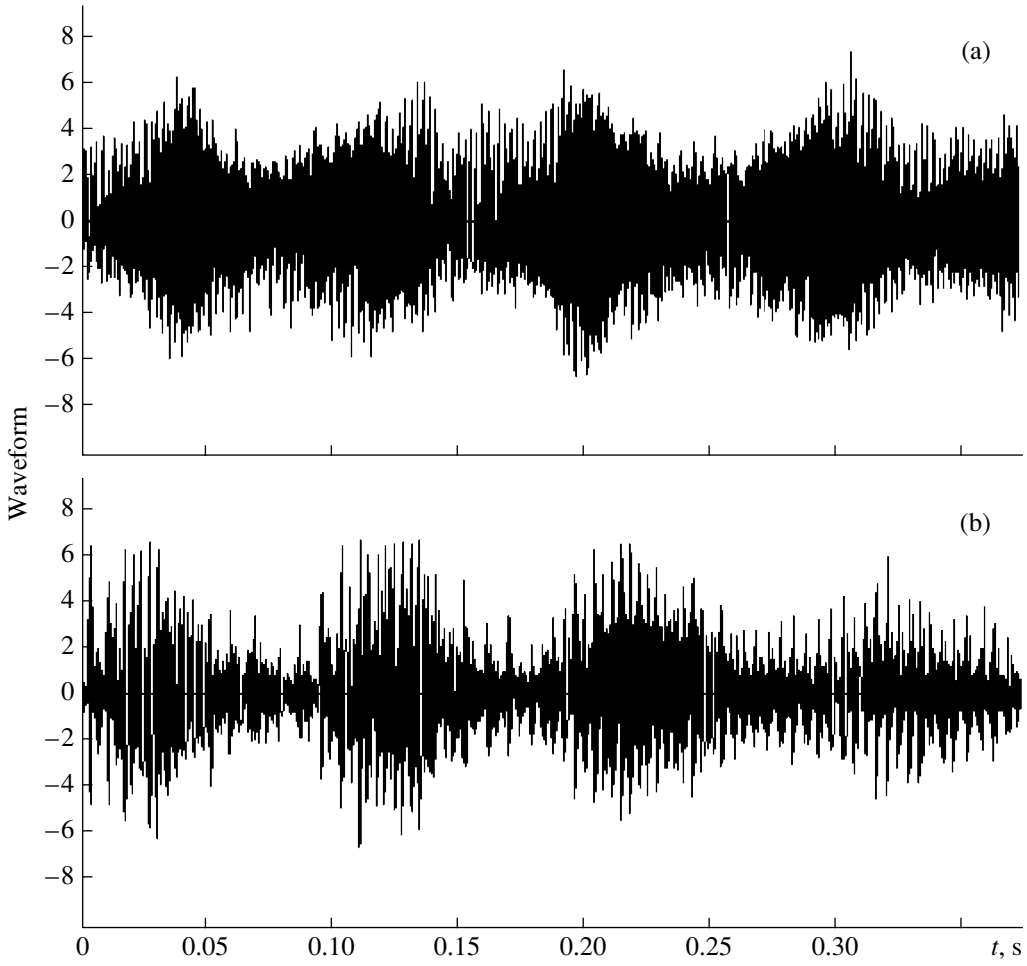


Fig. 3. Oscillograms of the (a) Russian vowel "a" and (b) "bleat" of a white whale.

corresponds to a frequency of 8.3–12.5 Hz. In this case, the depth of amplitude modulation is about 70%, and in the case of frequency modulation it is about 40%. Figure 4 presents different cases of amplitude and frequency modulation combinations in vowel-like sounds of white whales that are perceived as a "bleat" (the upper curve is the envelope of the main signal and the lower curve is the envelope of intensity). In the "voice pitch," the vowel-like components of the communicative signals of white whales lie approximately within the range of human voices and are clearly distinguishable audibly. The relatively high "voices" constituted 20%, the medium ones, 50%, and relatively low ones, 30% of all recorded signals. For comparison, we give the data (Table 1) on the occurrence of human voices of different pitch [8].

The measurements of the frequency range of "vocalization" demonstrated that the whole set of studied signals lies within the limits of 100–300 Hz, which confirms perceptive observations on the presence of this component within the range of human voices (70–500 Hz for speech and 60–1300 Hz for singing). In this case, perceptive division into low, medium, and high "voices" of white whales corresponded to grouping of detected signals within the intervals of 100–200 Hz, 125–250 Hz, and 150–300 Hz. For comparison, we present data (Table 2) measuring the frequency range of standard timbre–pitch types of human voices [8].

As one can see from the given data, some parts of the common range considerably overlap. It is also known that, in the situations connected with emotional excitation, the range of the fundamental frequency variation of both human speech and components of communicative signals of white whales can be extended.

As we have already noted, whistle components are also present in the complex of a communicative signal of a white whale (72%) apart from "vocal" components, and 35% of the whistle components are "clear" in their timbre (Fig. 1) while 37% are "noisy" (Fig. 5a). Whistle components occupy the frequency range of 4–18 kHz.

A "clear" whistle has a harmonic structure and occupies a wide frequency band. Oppositely, a "noisy" whistle is represented by narrowband noise. Figure 5b shows a spectral–time block including a short whistle

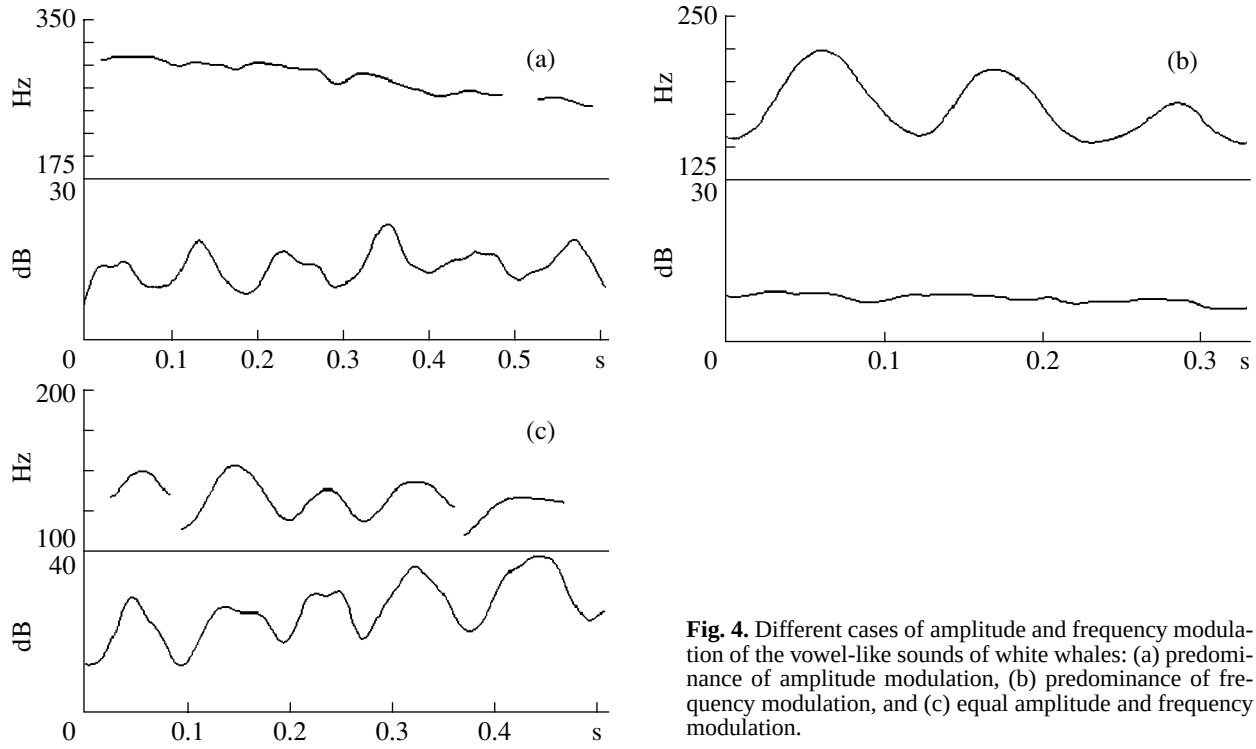


Fig. 4. Different cases of amplitude and frequency modulation of the vowel-like sounds of white whales: (a) predominance of amplitude modulation, (b) predominance of frequency modulation, and (c) equal amplitude and frequency modulation.

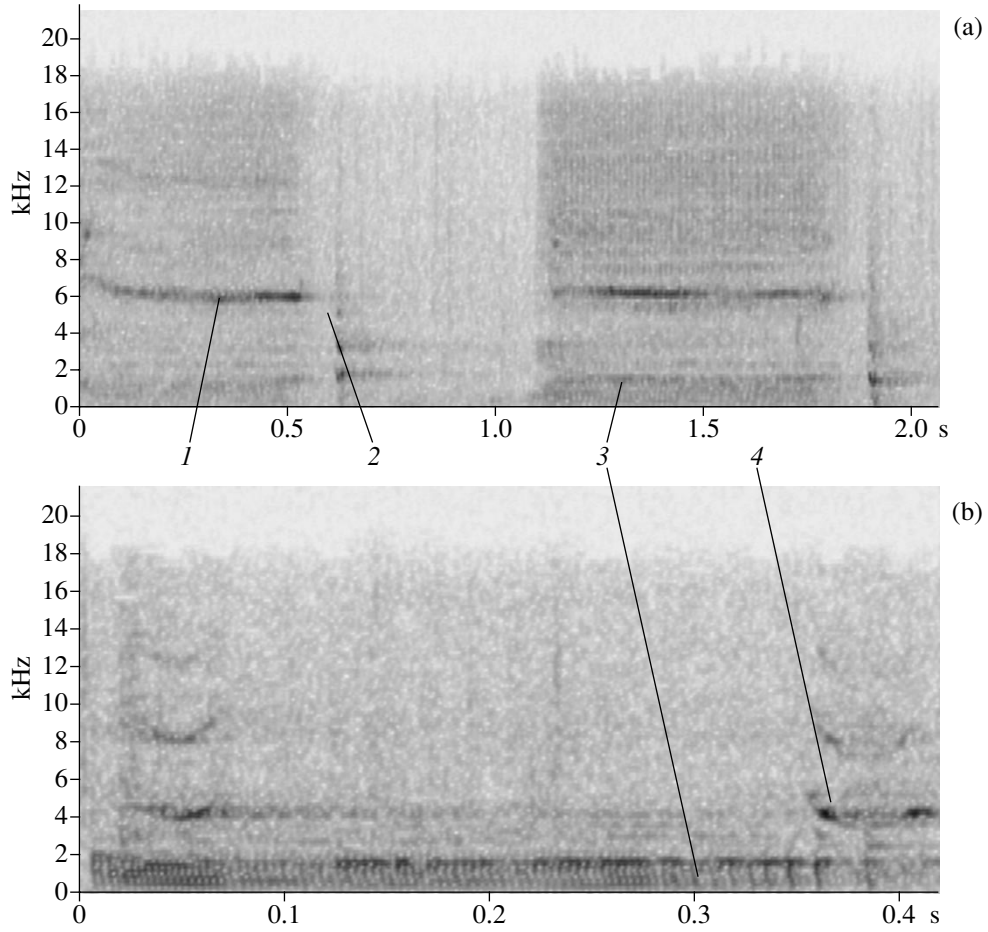


Fig. 5. Spectral–time blocks of communicative signals of a white whale, which include (*1*) a narrow-band "noisy" whistle, (*2*) a "resonant" stroke, (*3*) a "bleat," and (*4*) "chirping".

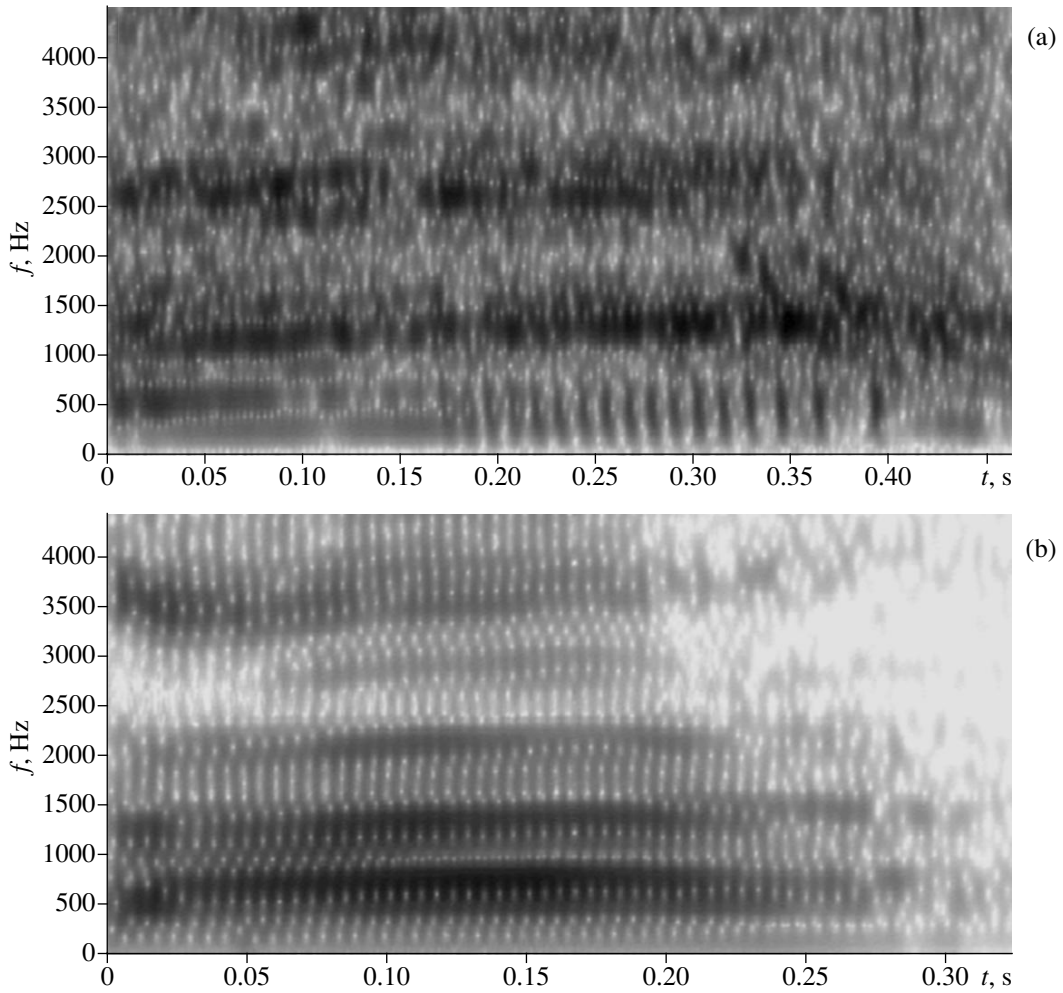


Fig. 6. (a) A sonogram of a vowel-like sound of a white whale: F1 = 500 Hz, F2 = 1100 Hz, F3 = 2700 Hz, and F4 = 4000 Hz. (b) A sonogram of a stressed vowel "a" of Russian speech in the word *vrata* (*gates*): F1 = 500 Hz, F2 = 1250 Hz, F3 = 2000 Hz, and F4 = 3500 Hz.

perceived as similar to a bird's "chirping." Blocks with such structure, where a vowel-like sound is "framed," occurred regularly in the observation data and were, probably, an individual feature of an animal.

Pulsed components, which can be seen in Figs. 1 and 5, occurred in 52% of the analyzed signals. According to their audible perception, they can be separated into four types:

(1) a click (a "dry" sound), 14%;

(2) a resonant stroke (a high tone is audible), 6%;

(3) a flat stroke (low noise), 4%;

(4) sounds perceived audibly as "ik," "ok," and "chmok" and representing a combination of short tonal and noise pulses, 28%.

The characteristic feature of the latter was that such a sound often concluded a "bleat" with a noisy whistle and the whistle transformed into this kind of sound.

The methods of searching for individual distinctive features of communicative signals of white whales were adopted from forensic phonoscopic inquiry [9], where the personality of a speaker can be identified by the totality of linguistic and acoustic characteristics of speech and voice. An oral statement is a spectral–time structure consisting of a series of components with certain frequency and time parameters, which characterize both the speech tract of a speaker and his individual manner of speaking. Under certain conditions, this provides an opportunity to compare a given statement with analogous statements pronounced by other persons or the same person in a different speech situation. These parameters are the average sound duration, the average frequency of the fundamental tone of speech, and the range of its change, as well as the average values of formant frequencies of stressed vowels, which integrally characterize the acoustic properties of the speech tract and the acoustic characteristics of speech of a person.

In this connection, the vowel-like components of the communicative signals of white whales are of the greatest interest, because they are differentiated more finely

in pitch and timbre of sound, as compared to, for example, whistle components lying in the frequency range where fine audible discrimination by pitch and timbre is difficult. Therefore the utilization of vowel-like sounds as a natural acoustic marker of white whales looks promising, since these sounds have a formant structure similar to the formant structure of speech vowels [10], which characterize the individual properties of the speech tract of a person.

Figure 6, for comparison, shows the sonograms of a vowel-like sound of a white whale and a speech vowel, where one can see the formant regions (F1, F2, F3, and F4 are the average frequencies of formants).

The duration of vowel-like sounds of white whales is approximately two to three times greater than that of the vowels of human speech and constitutes 400–700 ms. Longer (about 1.5–2 s) vowel-like sounds with a deep frequency and amplitude modulation of about 10 Hz, which are audibly perceived as a "bleat," belong to the same class of sounds.

Thus, to use the communicative signals of white whales as a natural acoustic marker, we took into account the duration of a vowel-like sound or "bleat"; the range and the average values of its fundamental frequency; the shape of the frequency profile; and the values of F1, F2, F3, and F4.

Grouping the detected communicative signals of white whales according to the totality of the indicated characteristics of vowel-like sounds allowed us to conditionally distinguish about 40 different animals. This number approximately coincided with the number of animals detected visually during the observation period. However, to obtain the full pattern characterizing an animal by the emitted communicative signal, it is necessary to take also into account the specific features of other components, i.e., the whistle and pulsed ones.

REFERENCES

1. V. M. Bel'kovich, E. N. Izekeeva, and V. S. Baranov, in *Proceedings of X Session of the Russian Acoustical Society* (GEOS, Moscow, 2000), p. 385.
2. V. M. Bel'kovich, R. A. Belikov, L. V. Zlatoustova, and S. A. Kreĭchi, in *Proceedings of XIII Session of the Russian Acoustical Society* (GEOS, Moscow, 2003), p. 265.
3. S. A. Kreĭchi, V. M. Bel'kovich, and R. A. Belikov, in *Proceedings of International Conference on Holarctic Marine Mammals* (Moscow, 2002), p. 144.
4. V. M. Bel'kovich and O. I. Kirillova, Zool. Zh. **79** (1), 89 (2002).
5. R. A. Belikov and V. M. Bel'kovich, in *Proceedings of XI Session of the Russian Acoustical Society* (GEOS, Moscow, 2001), p. 199.
6. R. A. Belikov and V. M. Bel'kovich, in *Acoustics of the Ocean: Proceedings of IX School–Seminar of Academician L. M. Brekhovskikh Combined with XII Session of the Russian Acoustical Society* (GEOS, Moscow, 2002), p. 179.
7. L. V. Zlatoustova, *Phonetic Units of Russian Speech* (Mosk. Gos. Univ., Moscow, 1981).
8. L. V. Zlatoustova and S. A. Kreĭchi, in *Proceedings of XIII Session of the Russian Acoustical Society* (GEOS, Moscow, 2003), p. 265.
9. E. I. Galyashina, *Forensic Phonoscopic Expertise* (Moscow, 2001).
10. L. V. Zlatoustova and S. A. Kreĭchi, in *Proceedings of III All-Russian Conference on the Theory and Practice of Speech Studies* (Moscow, 2003), p. 69.

Translated by M. Lyamshev

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 295–304. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 357–368.

BIOLOGICAL ACOUSTICS

What Do Evoked Potentials Tell Us About the Acoustic System of the Harbor Porpoise?

N. G. Bibikov

Andreev Acoustics Institute, Russian Academy of Sciences, ul. Shvernika 4, Moscow, 117036 Russia
e-mail: bibikov@akin.ru

Received October 2, 2003

Abstract—The evoked acoustic potentials of the brainstem (EAPB) were detected from the brain, the skull, and the surface of the head of the harbor porpoise (*Phocaena phocaena*). Experiments were performed at the Karadag biological station (Crimea). Clicks, noise, and tone bursts of different frequencies within 80–190 kHz were used as stimuli. The time and frequency selectivities of the auditory system were estimated by the simultaneous and direct forward masking methods. The minima of EAPB thresholds were usually observed in a frequency range of 120–140 kHz, which corresponded to the main spectral maximum of the species-specific echolocation signal. In addition to the regular EAPB, a pronounced off-EAPB was observed. In the aforementioned frequency range, a frequency selectivity (Q_{10} of about 10) was revealed by the direct forward masking method. The EAPB could be measured up to a frequency of 190 kHz, but outside this high-resolution region (outside the ultrasonic "fovea"), the frequency selectivity was weak. A simultaneous masking of a click by a tone was strong only when the delay of the click with respect to the masker onset was smaller than 1.0 ms. In a continuous regime, the tone (unlike noise) produced only a weak masking. The response to a small intensity increment of 1–4 dB was rather strong. In the frequency range of 120–140 kHz, this response exhibited a nonmonotone dependence on the signal level. The time resolving power, which was measured by the EAPB recovery functions for double clicks of various levels, was rather high, even when the intensity of the test signal was 18 dB lower than the masker level. Experimental data show that the auditory system of the harbor porpoise is tuned to detecting ultrasonic echo signals in the frequency range within 120–140 kHz. A hypothesis is put forward that the acoustic system of the harbor porpoise allows the animal, from analyzing echo signals, to estimate not only the distance to the target and the target's intrinsic properties but also the speed with which the target is approached, the latter estimate being presumably obtained on the basis of the Doppler effect. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The acoustic system of both humans and animals still remains an unsurpassed instrument for analyzing and classifying complex sound signals. The analyzing abilities of most vertebrates, even those with relatively nonspecialized sound processing systems, are unique. For scientists, of special interest are animals who not only are capable of analyzing various sounds, but also developed unique abilities to classify sound signals within a certain specific subclass. Examples of such animals are owls, who are capable of hunting at night by means of the passive location of very weak sound sources. Another example is *Homo sapiens*, for whom the ability for complex communication through sound signals proved to be one of the key factors of its vigorous progress within the last several thousand years. The attention of researchers is also attracted to animals using their own sound signals for orientation in space, i.e., echolocating animals. The full set of species possessing the echolocation ability remains unknown. Their most famous representatives are bats and toothed whales. Both these groups of animals are predators and use echolocation mainly for hunting, often instead of vision.

Understanding of the mechanism underlying the operation of any sensory system is impossible without studying its electrophysiological activity. To study the mechanism of echolocation, bats are especially convenient objects. The detailed investigation of the specialized mechanisms of echo-signal analysis by neurons in a bat remains one of the best achievements of research into central auditory mechanisms [1]. However, this does not mean that the other large group of animals, which use echolocation in other conditions (in water, at large distances from the source), does not deserve a similar close investigation.

Electrophysiological studies of the brain and the auditory system of dolphins encounter many difficulties associated with both ethical and technical problems. The large size of the animals, the difficulties in their housing, and their specific reactions to anesthesia have resulted in researchers almost giving up the attempts to study the firing activity of neurons in these animals. However, because the properties of the auditory system of dolphins attracted great interest, in the 1960s and 1970s the electrophysiological and morphological studies of dolphins have become a rapidly progressing area of research. Studies were carried out

1063-7710/04/5003-0295$26.00 © 2004 MAIK "Nauka/Interperiodica"

using electrodes inserted into the brain [2–10]. A publication that appeared in 1981 reported on the measurement of potentials by an electrode inserted into the skull of a dolphin without a trepanation [11]. In 1983, a small paper describing the potentials measured immediately on the head of a harbor porpoise was published [12]. Later, this method was considerably improved, mainly by the researchers from the Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences (the laboratory headed by A.Ya. Supin). Starting from 1985, the noninvasive method of measuring the brainstem potentials on the surface of the head of an animal has become dominant, primarily because of the appearance of more strict requirements concerning the ethical aspects of experiments with highly developed animals [13–22].

In spite of the numerous advantages of the noninvasive measuring technique, it has certain limitations. In fact, despite the considerable efforts, we still do not know the exact points of origin of the brainstem potential oscillations in most species studied, including humans (presumably, an exception is the first positive potential maximum, which is determined by synchronized pulses of the auditory nerve fiber). Most likely, this compound action response called evoked acoustic potential of the brainstem (EAPB) represents a superposition of many potentials. In clinical practice, at least before the wide acceptance of methods based on detecting the otoacoustic emission, the EAPB measurement served as the main method of objective audiometry providing fast diagnostics of the auditory system without feedback from the individual under examination.

The noninvasiveness of the method of detecting the EAPB from the head surface is very important for studying rare and precious animals, including various kinds of toothed whales. However, the method of local, or cranial, measurement of potentials has certain advantages from the viewpoint of studying the auditory processing of signals inside the brain of an animal: it provides greater signal amplitudes and a possibility to analyze the signal components in more detail.

Naturally, any measurement of intracranial brainstem potentials in both acute and chronical experiments was only possible for a small number of dolphins belonging to the most common species. In fact, such measurements were performed for only two species: the bottlenose dolphin (*Tursiops truncatus*) [2, 4, 6, 23] and the harbor porpoise (*Phocaena phocaena*) [3, 5, 7–10]. The latter proved to be a more convenient object because of its small size and availability (at the time of the cited experiments, this species was often found in fishnets by accident). Paradoxically, electrophysiological experiments were stimulated by the difficulties in long-term keeping of dolphins in tanks, which is necessary for behavioral experiments. All this made the harbor porpoise the more popular object of investigation.

The echolocating signals of harbor porpoise possess some specific features, which suggest the presence of a specialized sensory processing mechanism in the acoustic system of this animal. Unlike the signals of many other toothed whales, the echolocating signal of the harbor porpoise is a short pulse with a carrier frequency of about 130 kHz rather than a broadband click [25–28].

This paper analyzes the results of studying harbor porpoise by both the intracranial method (data taken from the cortex) and surface measurements (data taken from the skin surface on the head of the animal). The main attention is concentrated on the ability of the active hearing [29] of the harbor porpoise to tune to analyzing signals that belong to the frequency range near the carrier frequency of the echolocating signal of this animal. Experiments were performed in the laboratory headed by N.A. Dubrovsky, at the experimental base of the Karadag branch of the Institute of Southern Sea Biology, National Academy of Sciences of Ukraine.

Some of the results described below were partially presented in the literature [7, 12, 14, 16, 30].

EXPERIMENTAL TECHNIQUE

The basic description of the technique can be found in [14, 16], and in this section the experimental technique is only briefly outlined. An awake animal was placed in a relatively large tank and kept on the water surface by a net. The upper part of the head of the animal and the upper part of the dorsal fin were in air, while all other parts of the body, including the sound-perceiving structures (the auditory canal and the bones of the lower jaw) were in water.

The signal was produced by a piezoelectric transducer, which was fabricated at the Andreev Acoustics Institute and calibrated by a B&K hydrophone in the frequency range from 50 to 200 kHz. The signal level was measured in decibels relative to 2×10^{-5} Pa. The measurements of the local potential and the potentials from the dorsal surface of the cortex were performed using the technique that was developed in the laboratory of A.Ya. Supin (see [4, 23]) for detecting the local potentials of the auditory cortex. The potentials from the head surface were measured using a broad metal plate pressed to the surface by a rubber ring (as those conventionally used in recording electroencephalograms). All the potentials were accumulated after 500–3000 sequential independent presentations. In all cases of studying the masking, when it is imposed on the response to a test stimulus, the point-by-point subtraction of the response to the masker from the response to the masker-plus-signal pair was used [16, 19].

RESULTS

General Characteristic of Responses and Their Frequency Dependence

When the measurements are performed near the mean line of the head, 2–3 cm behind the bregma, the

form of the EAPB was almost the same for different sound signals. At a fixed signal, the fast components obtained from the cortex surface (away from the auditory cortex zones), the skull surface, and the head surface are almost identical. Figure 1 shows typical responses detected from the skull (Fig. 1a) and the head surface (Fig. 1b) under the stimulation by a tone signal with a frequency of 140 kHz, a sound pressure of 1 Pa, and a duration of 6 ms. The responses differ in their amplitudes; besides, in the case of the noninvasive measurement on the head surface, the initial response component is more pronounced (Fig. 1b). The slow response, which, presumably, is associated with the activity of the auditory cortex zones, is clearly noticeable only in the case of the intracranial measurement. It is important to note that the dependences of the amplitude and latency of the EAPB peaks on the signal level were of the same character for intracranial and surface EAPB measurements (see also Fig. 2 in [12]). Moreover, the measured EAPB was analogous to the local potential detected near the inferior colliculus [2, 10].

Since the main components of the EAPB of a dolphin (as well as a human) are positive in sign, we can classify the peaks according to their positive maxima. The main component that appears with a delay of about 3 ms can naturally be called PIV peak. Its amplitude measured from the maximum to the subsequent negative deviation will be used in our study (as in many other studies) as the measure of the EAPB value. Although the origin of this component may be rather complex, it mainly represents the synchronous discharge of the lateral lemniscus fiber, which connects the auditory centers of the medulla and the mid-brain. This is evidenced by numerous data of comparative physiology, as well as the results of the direct measurements near the lateral lemniscus of the harbor porpoise, where an intense and almost single-peak potential was detected with a latency period corresponding to the delay of the IV component of the EAPB (see Fig. 24 in [10]).

Since the hearing of dolphins mainly belongs the high-frequency range, the audiogram is usually studied on the basis of responses to very short (1–2 ms) but still narrowband signals. In Fig. 2, the circles show the dependences of the EAPB threshold levels on the frequency of tone bursts for two animals. According to our data, the harbor porpoise has a pronounced sensitivity maximum in the frequency band within 120–140 kHz, and this maximum is observed for all animals studied in our experiments. A similar maximum was observed for many objects in the local potential measurements [3, 5, 8–10]. Above 140 kHz, the high-frequency branch of the audiogram exhibits a steep rise, but a response to a signal of 190 kHz can still be detected.

Presence of Responses to Stimulus Termination

A characteristic feature of evoked brainstem potentials in harbor porpoise is the presence of a response to the stimulus termination (the off-EAPB). This effect has been described earlier in papers reporting on the measurement of local responses in practically all nuclei of the brainstem of harbor porpoise [9, 10]. The effect was observed with all methods of measurement (Fig. 1). The frequency dependences of a regular EAPB and an off-EAPB may differ substantially. The lowest off-EAPB threshold either corresponds to frequencies slightly exceeding the highest-sensitivity frequency of the EAPB (Fig. 2a) or depends little on frequency within 80–150 kHz (Fig. 2b). As a result, the difference between the thresholds of the regular EAPB and the off-EAPB usually reaches its maximal values in the frequency range of 120–130 kHz.

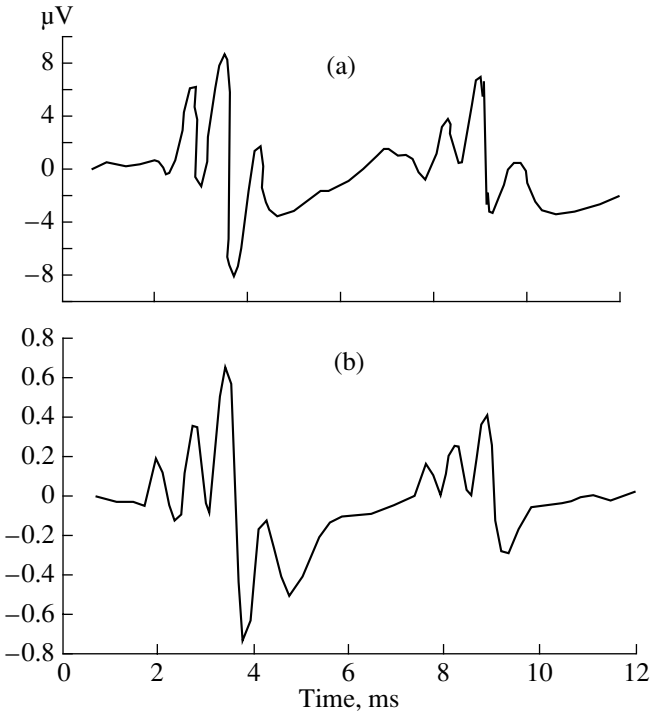


Fig. 1. Typical compound action potentials of the brainstem of the harbor porpoise in response to a tone burst with a frequency of 140 kHz, a level of 1 Pa above the threshold, and a duration of 6 ms: the active electrode is positioned (a) in the skull bone and (b) on the surface of the head. The positive polarity is directed upwards.

Figure 3 presents the dependences of the EAPB (Fig. 3a) and the off-EAPB (Fig. 3b) on the signal level. The curves are normalized with respect to the maximal EAPB amplitude at the corresponding frequency. The behavior of the amplitude dependence of the EAPB is approximately the same for all frequencies. However, this is not the case for the responses to the signal termination. One can notice the high amplitude of the off-EAPB at a frequency of 140 kHz and the weak dependence of the off-EAPB amplitude on the signal level when the signal belongs to the lower part of the frequency range under study (80–100 kHz). At frequencies of 100 and 90 kHz, in the superthreshold region,

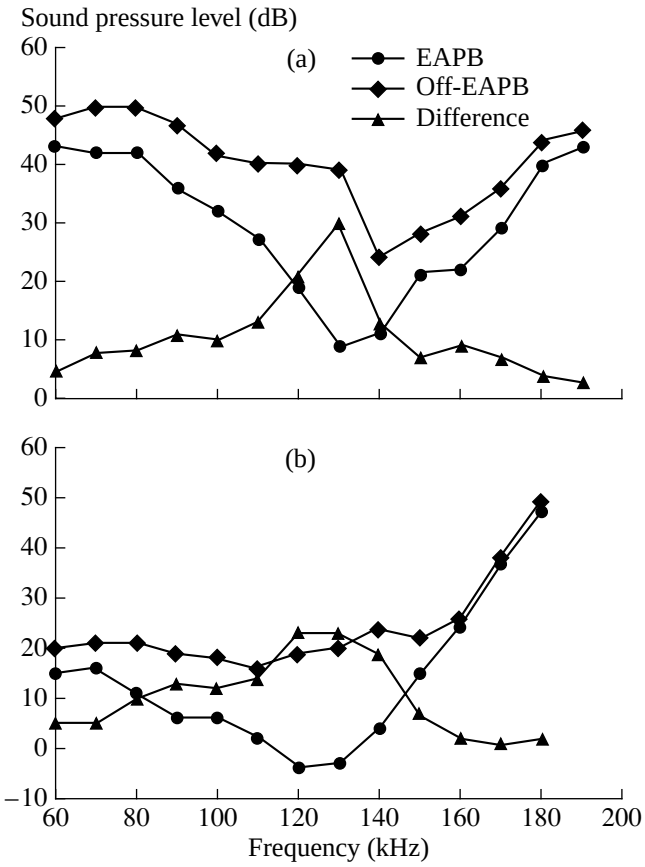


Fig. 2. Thresholds corresponding to the appearance of the compound action potentials of the brainstem in response to the onset of a tone burst (EAPB, circles) and to the termination of a tone burst (off-EAPB, diamonds) versus the frequency. The triangles show the difference between the thresholds in decibels. The data are obtained (a) from the head surface and (b) from the skull for two different animals.

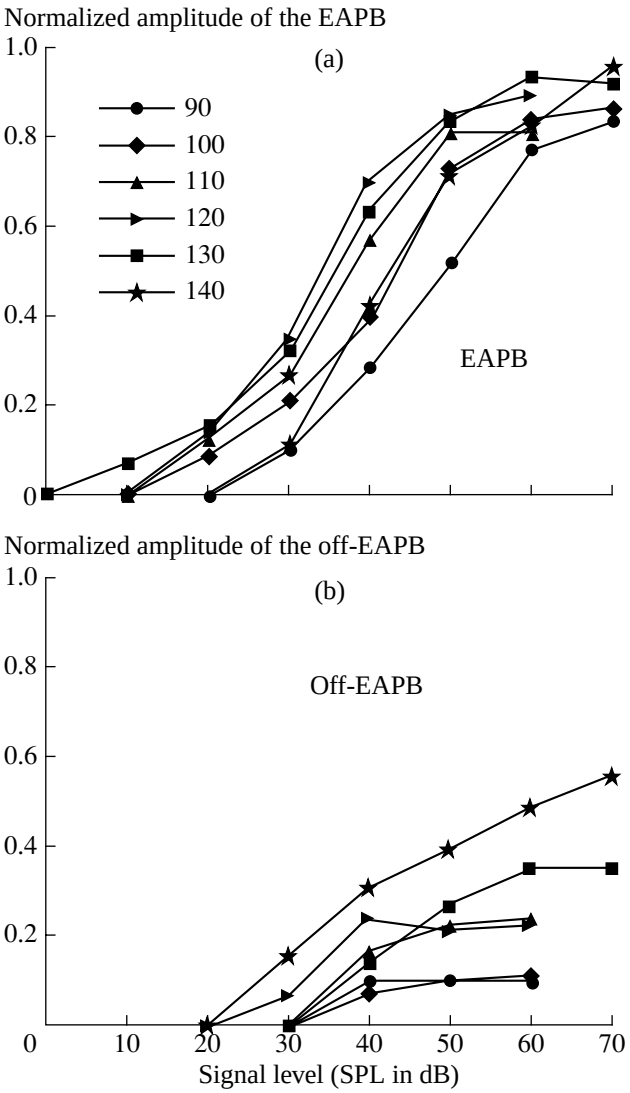


Fig. 3. Amplitude of (a) EAPB and (b) off-EAPB in response to tone bursts versus the signal level. In all cases, the data are normalized to the maximal EAPB amplitude at a given frequency. The frequencies used in the experiment are indicated in Fig. 3a.

the off-EAPB amplitude is practically independent of the signal level.

Dynamics of Simultaneous Masking

The special adjustment of harbor porpoise to analyzing the sounds in the frequency range of 120–140 kHz was demonstrated in a number of our experiments, in particular, in studying the simultaneous masking of a broadband click by masking tone bursts of different frequencies. Such a masking strongly depended on the time position of the click with respect to the tone burst onset. Figure 4 shows the time dependence of the relative amplitude of the response to a click on the click delay with respect to the masker onset. The maskers used in this case were 130- and 150-kHz tones and broadband noise. The response to an isolated signal corresponded to 100% in all cases. The relative amplitudes of the responses to individual maskers are shown by symbols on the ordinate axis.

These dependences are noticeably different for tone signals and noise. The tone signals cause a strong masking only when the click occurs immediately after the onset of the masker, while the noise masking develops more slowly. In the continuous regime (when the click occurs within more than 2–4 ms after the masker onset), the noise signal masks the click much more strongly than with the tone maskers.

A strong masking of the click by the initial part of a 130-kHz tone is observed. When the click is delayed by approximately 0.2 ms, the response to it is almost completely suppressed. Presumably, this testifies to the fact that, at the initial instant of the signal presentation, a very large population of neurons proves to be excited at

all peripheral levels of the auditory system of the dolphin.

In the presence of all maskers studied, especially in the presence of the 130-kHz tone, the effect of backward masking is observed. The response to the click is also suppressed when the masker is delayed by 0–0.5 ms with respect to the signal (not shown in Fig. 4).

Direct Sequential Tone-by-Tone Masking

One of the most used methods of estimating the frequency selectivity of the acoustic systems of different animals is the method of masking of one tone by another. Taking into account the weakness of the simultaneous masking under the effect of a continuous masker (Fig. 3) and also the literature data testifying to the efficiency of using a direct forward masking for the determination of the frequency selectivity [31], we used this method in our experiments. The duration of the masker usually was 3 or 5 ms, the duration of the test tone was 0.5–1 ms, and the interval between the masker and the signal was 1–2 ms. Because of the presence of the off-EAPB due to the masker, the subtraction procedure was used in all cases to separate the response to the test tone. Data of two such experiments are shown in Fig. 5. In one case (Figs. 5a–5c), both the frequency and level of the test tone were fixed. As the level of the masker was increased, the amplitude of the response decreased. This dependence was steepest when the frequencies of the masker and the signal were close to each other (Figs. 5a, 5b). The dependence of the masker level necessary for the 50% suppression of the response to the signal (Fig. 5c) is conventionally considered as the frequency tuning curve.

In another version of the experiment (Figs. 5d, 5e), the frequency and the level of the masker were fixed. In this case, an increase in the signal level caused an increase in the response. The latter increase was slowest for signal frequencies close to the masker frequencies. The dependence of the signal level necessary for obtaining a fixed amplitude of the response on the signal frequency (Fig. 5f) also reflects the frequency selectivity of the system.

In both cases, the frequency selectivity was detected at 130 kHz and proved to be almost identical (the values of Q_{10} were approximately equal to 10). The estimates of the frequency selectivity obtained by the two aforementioned methods were found to be almost fully coincident, despite the fact that the cumulative form of the audiogram should affect the selectivity estimated by these methods in opposite ways.

Response to an Increment

In several experiments, we studied the EAPB arising in response to relatively small intensity increments of tone bursts. The beginning of such an increment was delayed with respect to the signal onset by 2–3 ms, so that, in this case, the subtraction procedure was not necessary. The responses to increments were clearly pronounced even for increments of 1–2 dB. They occurred at all frequencies used in our experiments.

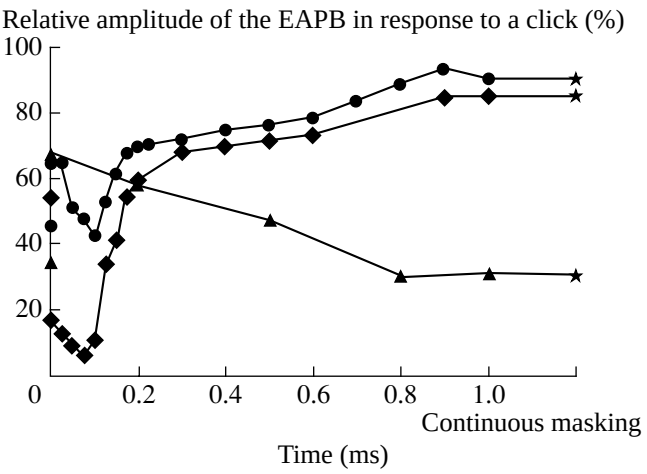


Fig. 4. Amplitude of the response to a click versus the interval between the masker onset and the click for three masker versions: 130-kHz tone bursts (diamonds), 150-kHz tone bursts (circles), and broadband noise bursts (triangles). The stars on the right indicate the amplitudes of the response to a click in the presence of a continuously acting masker, and the single dots on the left ordinate axis indicate the amplitudes of the response to the masker alone. All data are normalized to the amplitude of the response to a click in silence.

However, we have noticed a specific feature of these responses, which manifested itself at frequencies close to the frequency of the signal produced by the animal itself and, hence, to the minimum of the audiograms obtained by us. This feature consisted in a highly nonmonotone dependence of the response to the increment on the terminal signal level (Fig. 6a). The strongest response was observed at relatively low signal levels, and with an increase in the level, the response became weaker. At frequencies away from the optimum, although still lying in the far ultrasonic region, the aforementioned nonmonotonicity was either weakly pronounced or not observable (Fig. 6b).

Time Resolution

The time resolution of the acoustic system of dolphins has been much studied and discussed [7, 14, 16, 19]. In all our experiments, we observed with confidence the appearance of a response to the second of two clicks presented with an interval greater than 0.2 ms. In the case of a continuous sequence of clicks presented to the dolphin, the response was noticeable up to click rates of 2–3 kHz. These values are maximal for the EAPB of any animal and, presumably, exceed the corresponding values for other dolphin species. In addition to the known facts, we obtained data on the temporal recovery of the response to clicks of different levels.

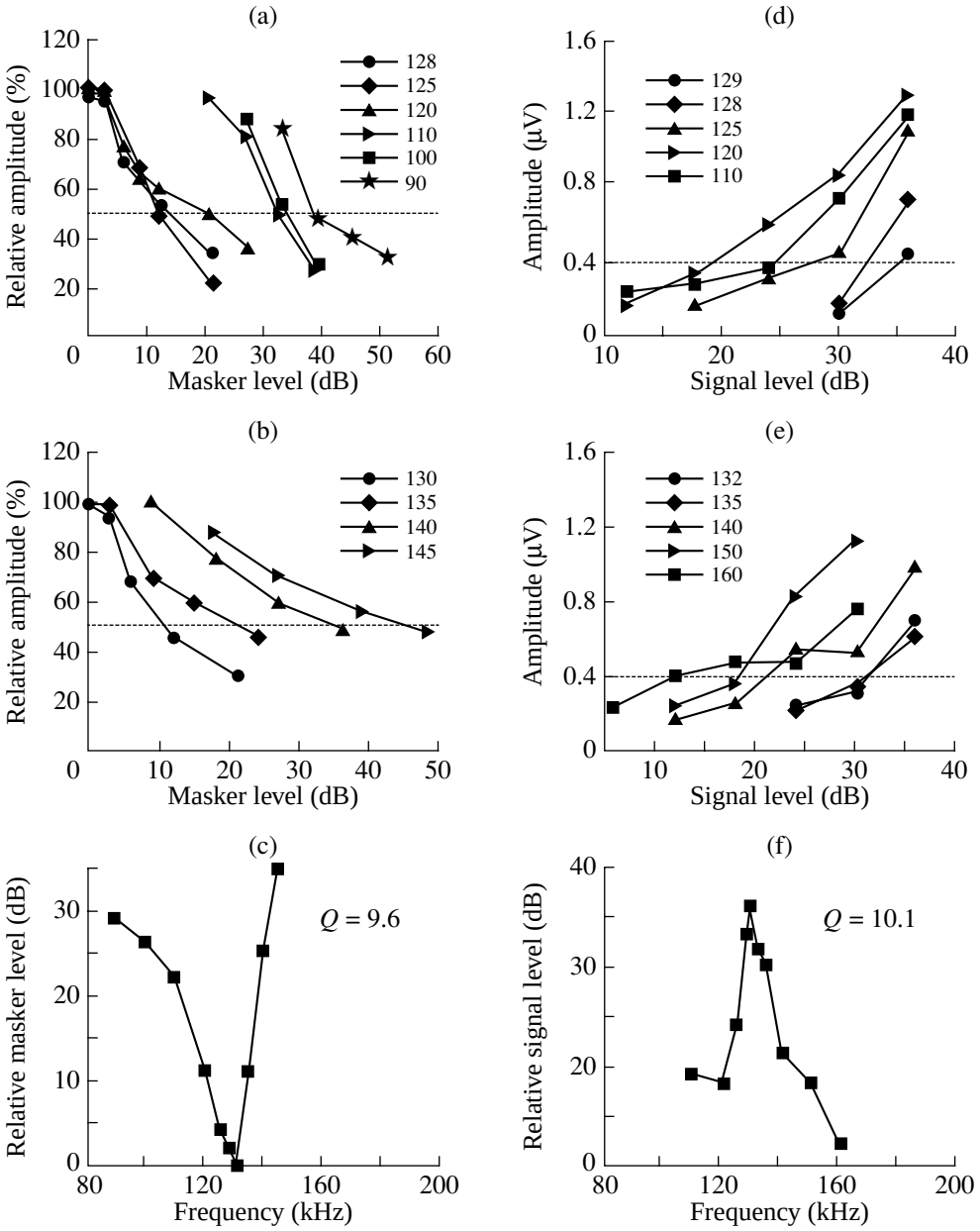


Fig. 5. Determination of the frequency selectivity by the method of direct forward masking (a–c) with a fixed signal and (d–f) with a fixed masker. (a, b) Amplitude of the response to a 130-kHz signal versus the masker level for different masker frequencies (indicated in the plots). The data are normalized to the amplitude of the response to the given signal in silence. (d, e) Amplitude of the response to signals of different frequencies versus the signal level in the presence of a 130-kHz masker (the amplitudes are given in microvolts). (c) The masker level producing a 50% suppression of the response to the signal versus the masker frequency and (f) the signal level producing a response with an amplitude of 2.5 μV versus the signal frequency. The values of Q_{10} are determined as the ratio of the central frequency (130 kHz) to the bandwidth at a level of 10 dB with respect to the extremum of the curve.

Figure 7 shows the curves characterizing the recovery of paired clicks, the first of which had an amplitude of 68 dB and the second had either the same amplitude (circles) or a level 12 (diamonds), 18 (upward triangles), or 24 dB (sideward triangles) lower than the first click level. All curves in this figure are normalized to the amplitude of the response to the second click presented in silence. A 12-dB decrease in the level of the second click has practically no effect on the recovery curve. Only a further decrease in the level of the delayed signal leads to a late recovery. However, even for a level difference of 18 dB, at a relatively large delay of 1.5 ms, we observed a recovery of the response almost identical with that observed for equal stimulus levels.

DISCUSSION

Audiogram

In connection with the specificity of the echolocating signal produced by harbor porpoise [25–28], the form of its audiogram is of special interest. All animals studied in our experiments exhibited the maximal sensitivity in the frequency band within 120–140 kHz. Evidently, the exact form of the audiogram may vary for different animals (Fig. 2). It may also depend on the accuracy of sound pressure calibration in the experimental tanks, which usually have limited dimensions. However, since the size of the reception zone of a dolphin (no matter whether it is the tympanum or the bones of the lower jaw [32]) should exceed the wavelength at frequencies higher than 100 kHz (1.5 cm), it seems unlikely that the interference in the tank could affect the results, especially in view of the fact that the scatter of the sound pressure values at a given point in the tank usually does not exceed several decibels. The number of electrophysiological studies testifying to the presence of a sensitivity maximum within 120–140 kHz for harbor porpoise is fairly large, each of these studies being performed with several animals [7–10, 14, 16]. However, in the literature, one also can find data that testify to a considerable scatter of audiograms for different specimens of the species.

In particular, a recent publication [24] characterized the surface EAPB of a single harbor porpoise specimen. The shape and amplitude of the response to the click correlated more closely with the responses detected by us on the skull surface rather than with the responses obtained in our experiment with a surface electrode. The dependences of the amplitude and latency of the PIV peak on the signal level agreed well with our data. However, no such agreement was obtained for the behavior of the threshold curve. The maximum of hearing sensitivity was observed for the aforementioned animal at a frequency of 54 kHz. This result noticeably differs from our data and from the results of other researchers: the frequency range of 50–60 kHz was found to corresponded to a pronounced minimum in the sensitivity determined by measuring the cortex responses of the same object (see Figs. 9 and 10 in [33]) and the behavioral audiogram [34]. Presumably, in some cases, harbor porpoise may exhibit considerable individual deviations. It is also possible that the animal studied in [24] had a high-frequency hearing loss, for example, because of cochlea disorders. The steep growth in the response amplitude with increasing stimulus level, which was observed by the authors of the cited publication, could also testify to the effect of loudness recruitment typical of many cochlear disorders. However, these authors, as have other researchers, noted that the hearing range of harbor porpoise is shifted to higher frequencies compared to, e.g., bottlenose dolphin.

A considerable scatter in the audiograms of harbor porpoise was also observed in another experiment [35], the author of which used the method of threshold determination by the galvanic skin response to sound (the response was supported by a weak electrical shock on the skin). As a typical example, the author presented an audiogram with a sharp minimum in the threshold curve at a frequency of 128 kHz.

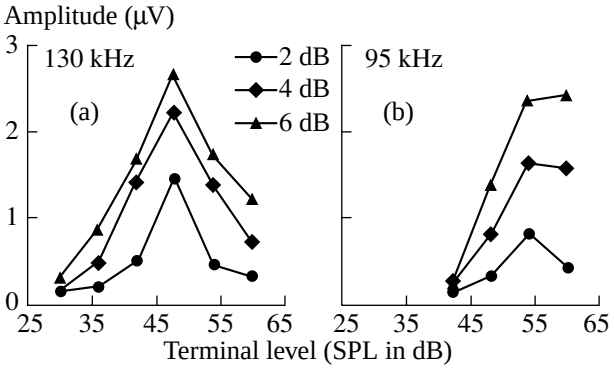


Fig. 6. Amplitude of the response to a signal increment versus the terminal amplitude value. The increment values are indicated in the plots. The data are obtained for carrier frequencies of (a) 130 and (b) 95 kHz.

In the recent behavioral study [34], the highest sensitivity of one of the animals under study was observed within 100–130 kHz without any pronounced sensitivity maxima in this frequency range.

Considering all audiographic data known for the harbor porpoise, one can make the following conclusions: (1) harbor porpoise seems to have the highest-frequency hearing among all animals studied by now, including both vertebrates and invertebrates; (2) most of the harbor porpoise specimens studied exhibited a sensitivity maximum at frequencies approximately corresponding to the frequency of their echolocation signal. The first conclusion alone makes the study of this animal of special interest. Indirect data in support of the second conclusion can be obtained from the analysis of other results of our studies.

Origin of the Off-Responses

The response of a single neuron to a tone pulse sometimes contains not only the response to the signal onset but also a pulsed response to the signal termination. However, conventional laboratory animals exhibit a small number of such response elements. The latter may also appear under the stimulation by relatively intense tone signals with short rise times [36], but in this case the responses are presumably determined by broadband transient processes. The compound action potentials of the brainstem of humans and common laboratory animals may also contain the off-components, which are also determined by the responses to transient processes [37] or, sometimes (usually in the presence of noise), by an abrupt termination of the neuron response. In the latter case, the potential usually has an inverted

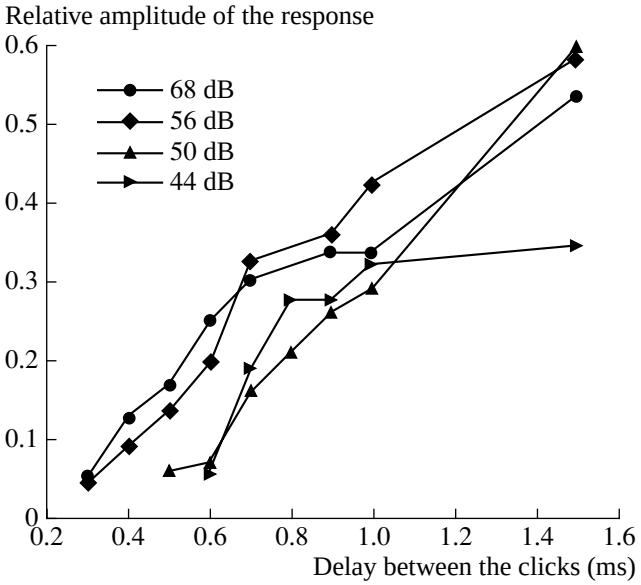


Fig. 7. Relative amplitude of the response to the second click of a pair of clicks versus the click spacing (in milliseconds). The amplitude of the first click is fixed at a level of 68 dB, and the amplitudes of the second click are indicated in the plot. All amplitudes are normalized to the response to the second click in silence.

form [38]. In the data under consideration, the latter mechanism is ruled out because of the closeness of the EAPB and off-EAPB shapes, as well as because of the expected weakness of the persisting response to high-frequency tone signals (Fig. 4). The response to transient processes usually appears under the effect of signals away from the optimal frequency. In our case, the contribution of such a component seems to be probable only in the presence of signals whose frequencies are far from the minimum of the audiogram. Such an origin of the off-EAPB arising under the action of relatively low-frequency tone bursts is evidenced by their small amplitudes even at high stimulus levels (Fig. 4). Naturally, at the end of a low-frequency signal, a small oscillation of the basal area of the cochlea must take place. However, under the action of high-frequency signals, the off-EAPB has a large amplitude and is similar to the initial EAPB in shape. For other dolphin species, the off-EAPB has not been a subject of special investigation (see also [9]). The analysis of the illustration, in which this potential is seen for a bottlenose dolphin (see, e.g., Fig. 13 in [13] and several figures in [20]), shows that, in this species, the off-EAPB is much less pronounced and its origin may be explained by an abrupt termination of the firing rate (the polarity of the main peak is negative).

The most prominent example of the off-responses in peripheral acoustic compound action potentials is given by a bat (*Pteronotus parnellii*) [39]. In this case, the shape of the potentials reproduces the shape of the initial EAPB, the thresholds of the potentials are close to each other, and the optimal sensitivity of the off-EAPB differs from the maximal sensitivity to the signal onset. Precisely the same situation is observed for the compound action potentials of the brainstem of the harbor porpoise.

The authors of [39] explain the presence of peripheral off-responses in the bat by the special adjustment of the basilar membrane area to the perception and fine frequency analysis of echo signals in the frequency range of 61–62 kHz (the spectral maximum of the species-specific echo signal). It seems reasonable to assume that a similar adjustment exists in the harbor porpoise, naturally, in the region of the basilar membrane tuned to the frequency of the fundamental component of the echolocation signal, which is about 120–140 kHz. One even may consider the possibility of individual distinctions and an adaptive tuning of the frequency of echolocation signals produced by harbor porpoise, by analogy with the mechanism presumably realized for bats (*Pteronotus parnelli*).

Frequency Selectivity

The frequency selectivity estimated by us using the method of direct forward masking of a tone by another tone at a frequency of 130 kHz proved to be approximately the same for the cases of signal or masker frequency variations and was approximately equal to 10 according to the Q_{10} criterion (Figs. 4a–4c). At other frequencies, the selectivity was much lower. In similar experiments on a cat [40], the values obtained for the quality factor were much lower (from 1.5 to 5.6) in the frequency range up to 10 kHz. Similar quality factors were obtained by us with the use of a test frequency of 90 kHz (Fig. 9 in [16]). Our results can be tentatively compared with the data obtained by the same method of direct forward masking of tone by tone for a bottlenose dolphin [22]. The authors of [22] used only several fixed masker and signal levels and very short durations of both signal and masker. Therefore, the quantity Q_{10} could not be determined quantitatively. However, at the frequency of 60 kHz used by the authors, the frequency selectivity measured by this method seems to be rather low.

Completely different results were obtained by the same authors in estimating the frequency selectivity by the methods of masking the responses to tones by a noise signal with a spectral dip. In this case, at a frequency of 90 kHz, the value of Q_{10} approached 20 [20]. A high selectivity was also obtained using the method of a simultaneous tone-by-tone masking [18].

Today, it seems that the estimate of the frequency selectivity by the method of direct forward masking of tone by tone is not optimal for such a specially adjusted animal as the harbor porpoise, whose neurons are likely to respond to a tone burst by only a short initial discharge and (possibly) a discharge at the termination of the signal. This is confirmed by our data on the dynamics of the masking of clicks by tone signals. When a

click is delayed with respect to the tone burst onset by less than 1 ms, the masking of the click by a high-frequency tone decreased several times, as compared to the case of delays of 0.1–0.2 ms (Fig. 4). One can assume that, in the experiments illustrated in Fig. 5, we virtually measured the masking of the on-response to the test tone by the off-response to the masker. It is of interest that, even in this case, the frequency selectivity at a frequency of 130 kHz was sufficiently high and, in any case, considerably exceeded the selectivity obtained at other frequencies. Note that the masking of a click by a noise signal was practically independent of the delay (see Fig. 2 in [16] and Fig. 4 of the present paper), so that the noise was an effective masker during the whole time of its presentation.

Time Resolution

The problem of time resolution is considered in most publications concerned with the analysis of electric activity evoked in the auditory system of dolphins by sound signals [2, 4, 6, 11–16, 19, 21, 22, 41]. In time resolution, dolphins outperform virtually all other animals by an order of magnitude. A noticeable recovery of the response is already observed when the intervals between clicks are about 0.2–0.3 ms. The recovery of a sequence of clicks is observed up to a click rate of 2000–3000 pulse/s [16, 41]. That high time resolution facilitates the analysis of the fine structure of echo signals and, in particular, the discrimination of the intervals between direct and delayed echo signals (the typical values of such intervals are precisely of the aforementioned order of magnitude) [27, 42]. The capability of dolphins to perform a temporal analysis at intervals of 0.2–0.3 ms fully agrees with the hypothesis put forward by N.A. Dubrovsky, which consists in that dolphins possesses two different hearing mechanisms: one of them functions within the critical interval, and the other, outside it [29].

The possibility of such a fine temporal analysis by an inertial system (remember that the duration of one neural pulse exceeds the critical interval, even if the refractoriness property is neglected) was specially investigated in a number of model studies [14, 43–45]. The main idea of these studies was that, because of the presence of the internal noise component and the refractory period, part of the auditory nerve fibers prove to be incapable of responding to the first click but respond to the second click (either because of the change in the instantaneous noise level or because of the release from refractoriness). Since the number of auditory nerve fibers and neurons of other auditory formations of dolphins is rather great [46, 47], the absolute number of neural elements responding to the second signal proves to be substantial.

Our analysis of the responses to paired clicks of different amplitudes (Fig. 7) shows that even a weak echo signal with a very small delay relative to the transmitted signal can be analyzed almost without any interfering effect of the main echolocating stimulus. Note that a rapid recovery of the response to weak short delayed tone bursts was also observed in a bottlenose dolphin [22].

CONCLUSIONS

This paper summarizes some results obtained from the electrophysiological studies of the acoustic system of the harbor porpoise in 1976–1980, on the experimental base of the Karadag branch of the Institute of Southern Sea Biology, Academy of Sciences of the Ukrainian SSR. The purpose of the paper was the analysis of the data that point to the special adjustment of the aforementioned species to analyzing the signals with frequencies corresponding to the spectral maximum of the transmitted echolocation signals. We proceeded from the hypothesis (which is still not proved) that the harbor porpoise is capable of using echo signals not only for determining the distance to the target (from the time delay between the locating signal and the echo signal) and for analyzing the structure of the target (by analyzing the intervals between primary and secondary echoes), but also for determining the speed with which the target is approached (from the Doppler frequency shift of the echo signal). Although the indirect data collected in this paper testify to the ability of the object under study to use the latter mechanism, they do not allow one to state with confidence that this mechanism exists in dolphins. Now, it has become clear that some of the experiments were planned and performed not quite correctly from the viewpoint of solving the problem under consideration. For example, one should carry out experiments to study the ability of harbor porpoise to differentiate the signal frequency in the region of interest. Here, it only can be noted that the data obtained by measuring the galvanic skin response testify to a very high frequency sensitivity of the harbor porpoise [48].

ACKNOWLEDGMENTS

Experiments that provided the basis for this paper were carried out with the participation of A.V. Zanin, E.G. Vodyanaya, L.I. Mednikova, and S.V. Kosterin. I am deeply grateful to N.A. Dubrovsky for his assistance in organizing the experimental studies and for fruitful discussions. The work on this paper was supported in part by the Russian Foundation for Basic Research, project no. 02-04-48236.

REFERENCES

1. N. Suga, Neural Networks **3**, 3 (1990).
2. T. H. Bullock, A. D. Grinnel, E. Ikezono, *et al.*, Z. Vgl. Physiol. **59**, 117 (1968).
3. T. F. Ladygina, Zh. Évol. Biokhim. Fiziol. **10**, 48 (1974).
4. V. V. Popov and A. Ya. Supin, Fiziol. Zh. **62**, 550 (1976).
5. V. A. Voronov and I. M. Stosman, Zh. Évol. Biokhim. Fiziol. **13**, 719 (1977).

6. A. Ya. Supin, L. M. Mukhametov, T. F. Ladygina, *et al.*, *An Electrophysiological Study of Dolphin's Brain* (Nauka, Moscow, 1978).
7. A. V. Zanin and N. G. Bibikov, in *Problems of Neurocybernetics* (Rostov-on-Don, 1980), pp. 65–66.
8. V. A. Voronov and I. M. Stosman, Zh. Évol. Biokhim. Fiziol. **83**, 578 (1982).
9. V. A. Voronov, Fiziol. Zh. **68**, 1294 (1982).
10. V. A. Voronov and I. M. Stosman, in *Electrophysiology of the Sensory Systems of Marine Mammals* (Nauka, Moscow, 1985), pp. 35–56.
11. S. H. Ridgway, T. H. Bullock, D. A. Carder, *et al.*, Proc. Natl. Acad. Sci. USA **78**, 1943 (1981).
12. N. G. Bibikov, E. G. Vodyanaya, A. V. Zanin, *et al.*, in *Proceedings of X Acoustical Conference* (1983), p. 38.
13. V. V. Popov and A. Ya. Supin, in *Electrophysiology of the Sensory Systems of Marine Mammals* (Nauka, Moscow, 1985), pp. 85–106.
14. N. G. Bibikov, L. K. Rimskaya-Korsakova, A. V. Zanin, and N. A. Dubrovsky, in *Electrophysiology of the Sensory Systems of Marine Mammals* (Nauka, Moscow, 1985), pp. 56–84.
15. V. V. Popov and A. Y. Supin, J. Comp. Physiol. A **166**, 385 (1990).
16. N. G. Bibikov, in *Marine Mammal Sensory Systems*, Ed. by J. A. Thomas, R. A. Kastelein, and A. Y. Supin (Plenum, New York, 1992), pp. 197–211.
17. V. V. Popov, A. Y. Supin, and V. O. Klishin, in *Marine Mammal Sensory Systems*, Ed. by J. A. Thomas, R. A. Kastelein, and A. Y. Supin (Plenum, New York, 1992), pp. 269–276.
18. A. Y. Supin, V. V. Popov, and V. O. Klishin, J. Comp. Physiol. A **173**, 649 (1993).
19. V. V. Popov and A. Y. Supin, J. Acoust. Soc. Am. **97**, 2586 (1995).
20. V. V. Popov, A. Y. Supin, and V. O. Klishin, J. Acoust. Soc. Am. **102**, 3795 (1997).
21. V. V. Popov and A. Y. Supin, J. Comp. Physiol. A **183**, 519 (1998).
22. V. V. Popov, A. Y. Supin, and V. O. Klishin, J. Acoust. Soc. Am. **110**, 2227 (2001).
23. V. V. Popov and A. Ya. Supin, *Marine Mammals: Results and Methods of Investigation* (Nauka, Moscow, 1978), pp. 78–89.
24. V. O. Klishin and V. V. Popov, Dokl. Akad. Nauk **370**, 413 (2000).
25. N. A. Dubrovsky, P. S. Krasnov, and A. A. Titov, Akust. Zh. **16**, 444 (1971) [Sov. Phys. Acoust. **16**, 374 (1971)].
26. B. Mohl and S. Anderson, J. Acoust. Soc. Am. **54**, 1368 (1973).
27. V. M. Bel'kovich and N. A. Dubrovsky, *Sensory Systems of Orientation in Cetaceans* (Nauka, Leningrad, 1976).
28. W. W. Au, R. A. Kastelein, T. Rippe, and N. M. Schooneman, J. Acoust. Soc. Am. **106**, 3699 (1999).
29. N. A. Dubrovsky, in *Sensory Abilities of Cetaceans*, Ed. by J. A. Thomas and R. A. Kastelein (Plenum, New York, 1990), pp. 233–254.
30. N. G. Bibikov, J. Acoust. Soc. Am. **111**, 2371 (2002).
31. B. C. Moore and B. R. Glasberg, J. Acoust. Soc. Am. **80**, 93 (1986).
32. B. Mohl, W. W. Au, J. Pawloski, and P. E. Nachtigall, J. Acoust. Soc. Am. **105**, 3421 (1999).
33. V. V. Popov, T. F. Ladygina, and A. Ya. Supin, in *Electrophysiology of Sensory Systems of Marine Mammals* (Nauka, Moscow, 1985), pp. 5–35.
34. R. A. Kastelein, P. Bunskoek, M. Hagedoorn, *et al.*, J. Acoust. Soc. Am. **112**, 334 (2002).
35. M. N. Sukhoruchenko, Tr. Akust. Inst. Akad. Nauk SSSR **17**, 54 (1971).
36. E. A. Radionova, *Functional Characteristic of Neurons of the Cochlear Nucleus and the Auditory Function* (Nauka, Leningrad, 1971).
37. K. R. Henry, Hear. Res. **18**, 245 (1985).
38. R. D. Brinkmann and M. Scherg, Scand. Audiol. **8**, 27 (1979).
39. N. Suga, J. A. Simmons, and P. H. Jen, J. Exp. Biol. **63**, 161 (1975).
40. M. P. Gorga, J. McGee, E. J. Walsh, *et al.*, J. Acoust. Soc. Am. **73**, 256 (1983).
41. W. F. Dolphin, W. W. Au, and P. E. Nachtigall, J. Comp. Physiol. A **177**, 235 (1995).
42. V. A. Vel'min and N. A. Dubrovsky, Dokl. Akad. Nauk SSSR **225**, 470 (1975).
43. N. G. Bibikov, N. A. Dubrovskiy, G. A. Ivanitskiy, *et al.*, in *Proceedings of XI International Congress on Phonetic Sciences* (Tallinn, 1987), Vol. 3, p. 67.
44. L. K. Rimskaya-Korsakova and N. A. Dubrovsky, Sens. Sist. **4**, 92 (1989).
45. N. A. Dubrovskiy and L. K. Rimskaia-Korsakova, in *Marine Mammal Sensory Systems*, Ed. by J. A. Thomas, R. A. Kastelein, and A. Y. Supin (Plenum, New York, 1992), pp. 223–233.
46. G. F. Gao and K. Y. Zhou, in *Marine Mammal Sensory Systems*, Ed. by J. A. Thomas, R. A. Kastelein, and A. Y. Supin (Plenum, New York, 1992), pp. 39–52.
47. K. Osen and J. Jansen, J. Comp. Neurol. **125**, 223 (1965).
48. M. N. Sukhoruchenko, Fiziol. Zh. SSSR im. I. M. Sechenova **9**, 1205 (1973).

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 305–317. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 369–383.

BIOLOGICAL ACOUSTICS

Echolocation System of the Bottlenose Dolphin

N. A. Dubrovsky

Andreev Acoustics Institute, Russian Academy of Sciences, ul. Shvernika 4, Moscow, 117036 Russia
e-mail: dubrov@akin.ru

Received October 12, 2003

Abstract—The hypothesis put forward by Vel'min and Dubrovsky [1] is discussed. The hypothesis suggests that bottlenose dolphins possess two functionally separate auditory subsystems: one of them serves for analyzing extraneous sounds, as in nonecholocating terrestrial animals, and the other performs the analysis of echoes caused by the echolocation clicks of the animal itself. The first subsystem is called passive hearing, and the second, active hearing. The results of experimental studies of dolphin's echolocation system are discussed to confirm the proposed hypothesis. For the active hearing of dolphins, the notion of a critical interval is considered as the interval of time within which the formation of a merged auditory image of the echolocation object is formed when all echo highlights of the echo from this object fall within the critical interval. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Acoustic signals play a key part in the life of dolphins. Since light waves rapidly decay in water, distant perception for dolphins under the water surface is possible only by hearing. The purpose of the organ of hearing as a biological analyzer consists in the determination of the properties of a sound source. Since dolphins are echolocating animals, their organ of hearing can solve this problem in the passive mode, i.e., when the echolocation objects (e.g., fish) are sources of sound by themselves (primary sources), and in the active mode, when the objects are insonified by the dolphin's own echolocation clicks and produce echoes (secondary sources). In the first case, the organ of hearing operates as an independent system analyzing acoustic signals. In the second case, the organ of hearing must operate in combination with the dolphin's own sound source. In this case, it is appropriate to consider a specific complex analyzer, which includes the source of sound, the receiver, and the system of their interaction. This complex analyzer may be called the echolocation system.

The flow chart of the echolocation analysis is shown in a general form in Fig. 1. An acoustic click produced by (*1*) the source is focused by (*2*) a field-forming system and transmitted into (*3*) water. The echo signal (together with (*7*) noise and reverberation) produced by (*4*) the object under the effect of the echolocation click is again transmitted through (*3*) water and arrives at (*5*) the system of the auditory signal formation and then to (*6*) the system of formation and analysis of auditory images. The latter, being the most important and complex element of the whole echolocation process, also controls the adaptive process of generation and formation of the echolocation click (see the connection of unit *6* with units *1* and *2* in Fig. 1). Note also the possible connection between units *1* and *5*, which makes it possible to use delayed replicas of the probing pulse in the auditory analysis of echo signals, i.e., to perform the signal processing matched with the probing pulse.

The inclusion of the propagation medium, secondary sources, and sources of noise and reverberation into the general flow chart of the echolocation analysis is rather important, because the echolocation analysis adapts to both the acoustic characteristics of the medium and the parameters of secondary sources and interference.

The conditions of the operation of the hearing organ in the two aforementioned modes are different. The sounds from the primary and secondary sources (echo signals) differ in the volume of *a priori* data. In the first case, the animal does not know the direction of the arriving sound signal, its intensity, the distance to the

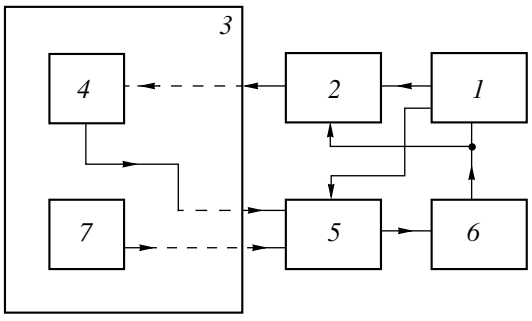


Fig. 1. Functional flow chart of echolocation analysis in bottlenose dolphins: (*1*) source of echolocation clicks, (*2*) system providing the formation of echolocation field, (*3*) echolocation data transmission channel (water medium and interfaces), (*4*) object of echolocation analysis, (*5*) system of auditory signal formation, (*6*) system providing the formation and analysis of auditory images and controlling the generation and formation of the echolocation field, and (*7*) sources of noise and reverberation.

1063-7710/04/5003-0305$26.00 © 2004 MAIK "Nauka/Interperiodica"

source, the instant of signal arrival, and the time–frequency characteristics of the arriving signal. Therefore, to provide an efficient perception of sounds from primary sources, the organ of hearing must possess a "panoramic" property; i.e., it must be always ready to receiving sound signals in all possible ranges of directions, distances, intensities, instants of arrival, and time–frequency characteristics. In the second case, the direction of the echo signal arrival is usually known, because it is opposite to the radiation direction. The distance to the secondary source, the intensity of the echo signal, and the instant of its arrival are only unknown at the very beginning of the echolocation. After the arrival of the first echo signals, these parameters become known. The frequency band of the echo signals usually coincides with that of the echolocation click and, hence, is also known to the animal. Thus, the main indeterminacy in the properties of the echo signal is concerned with its fine time–frequency structure.

Since the aforementioned differences in the operating conditions of the hearing organ in the active and passive modes are substantial, it is natural to assume that, in the course of the evolution of dolphins (and, possibly, other echolocating animals), two functionally specific subsystems of auditory perception could be developed. For brevity, we call them passive and active hearing [1–5]. For echolocating cetaceans, this assumption is even more justified, because their echolocation ability has appeared in the course of the secondary adaptation of their terrestrial ancestors, who had a fully developed passive hearing, to underwater conditions.

Different conditions of operation of passive and active hearing (i.e., hearing operating as part of echolocation system) should also cause some important distinctions in their characteristics. The necessity to perceive a signal arriving from a preliminarily unknown direction requires that the passive hearing be omnidirectional. Since the directional pattern is mainly determined by the ratio of the sound wavelength λ to interaural base d (or, in the more general case, to the characteristic size of the head or body of the animal), this requirement can be satisfied for the frequency band whose upper boundary $f_{max} = c/\lambda_{min}$ (c is the velocity of sound in water and λ_{min} is the minimal wavelength) is determined from the condition $\lambda_{min} < d$. Since the base d of an adult bottlenose dolphin is approximately equal to 15–20 cm, one obtains $f_{max} < (7.5–10.0)$ kHz; i.e., the passive hearing should mostly be low-frequency. This immediately provides long-range detection and identification of the sources of extraneous sounds, because, at the aforementioned frequencies, the attenuation in water is relatively weak. Ecologically, the most important biogenic sources of sound for bottlenose dolphins are the low-frequency sounds produced by fish, which also predetermines the relatively low frequency band of passive hearing.

By contrast, for the perception of secondary sound sources in the presence of noise and reverberation, a high directivity of radiation and reception is necessary, which can be achieved by increasing the frequency of radiation. Since the forward radiation direction is most expedient, the characteristic size determining the directivity of dolphin's sound source coincides with the size of a dolphin's head ($d \sim 20$ cm). Then, the central frequency characterizing the frequency band that can be efficiently used by a bottlenose dolphin for producing directional sound radiation should be determined from the relation $\lambda \ll d$, or $f \gg c/d$. This can be achieved at $f \sim 10c/d$, which yields $f \sim 75$ kHz. In this case, the main lobe width of the directivity pattern ($\alpha \sim 1.2\lambda/d$) is $\alpha \sim 7°$. The same characteristic spacing between two receivers (~20 cm) provides a highly directional reception at frequencies of the order of 100 kHz. Thus, the above consideration of the conditions of determining the direction to the source of sound suggests that the frequency band of passive hearing should lie mainly in the low-frequency region (about 1–10 kHz) and the frequency band of active hearing, in the high-frequency region (about 100 kHz).

For dolphins, the determination of the distance to primary sound sources is possible only with an *a priori* knowledge of the intensity of the source and the characteristic features of the time–frequency structure of the signal under different acoustic conditions [6]. To estimate the distance to secondary sources, it is necessary to measure the time interval between the instant of emission of the echolocation click and the instant of the echo signal arrival. Hence, the signal processing required for determining the distance to a source should be different for a source of signal and a source of echo signal, i.e., for passive and active hearing.

From the point of view of identifying primary and secondary sources, the tasks of passive and active hearing are also different. In the first case, there is a practically infinite variety of sounds that differ in their time–frequency and intensity structures, while in the second case, the frequency band of echo signals usually coincides with that of the echolocation click. In their time structure, echo signals from many objects have the form of a sequence of pulses similar in shape to the echolocation click but time-delayed with respect to each other.

Since the maximal dimensions of the food objects of a bottlenose dolphin are close to 20 cm, the time delays between the echo highlights should be smaller than 500 μs. In this interval of delay values, the auditory analysis of echo signals should be especially intricate for determining the properties of the object and for its classification and identification. In this connection, one should expect that active hearing possesses specific mechanisms for analyzing echo pulses whose durations do not exceed 500 μs.

Passive and active hearings also differ in the way of overcoming the adverse effect of noise and interference. In the passive case, the predominant type of noise is a low-frequency additive one, including biogenic

noise. In the active case, reverberation predominates. To overcome its effect, it is necessary to use short broadband echolocation clicks well resolved in space (time).

Unlike passive hearing, active hearing must function in close interaction with the source of echo signal and the medium, i.e., as an integral part of the echolocation system.

The estimates presented above confirm the validity of the hypothesis that dolphins possess a special auditory subsystem for analyzing secondary sounds (active hearing). By the time this hypothesis was published [1], some experimental data were obtained that evidenced in favor of the functional specificity of passive and active hearings. The most significant data were as follows: (i) matching the spectra of dolphin's echolocation clicks with the characteristic of the high-frequency absolute hearing sensitivity in the course of evolution [1]; (ii) presence of low differential thresholds in frequency below the frequency band of maximal hearing sensitivity, as in the case of nonecholocating mammals and humans [7]; (iii) selectivity of evoked electric potentials to short high-frequency signals with steep leading edges at inferior colliculus [8]; (iv) ability of dolphins to simultaneously emit (at least) two signals that differ in purpose and frequency band (namely, an echolocation signal in the higher frequency band and a communication signal in the lower frequency band), which, from the point of view of noise immunity, is expedient only when a simultaneous and parallel perception of these signals (i.e., echo signals and communication signals from other dolphins) is possible [9]; and (v) termination of the frequency differentiation with a change from continuous tone signals to pulses, which testifies to a fundamental difference in the perception of long signals and short pulses [10].

By now, the validity of the duplex auditory perception hypothesis for dolphins is supported by ample experimental data obtained by us and also by other research groups. The discussion of this data is the main purpose of the present paper.

Here, it is important to note the following fact: our experience showed that active hearing can be "activated" not only when the dolphin emits its own echolocation clicks but also when the animal, without producing such clicks, perceives external acoustic signals that are similar to echo signals in their characteristics: duration, frequency band, pulse repetition rate, and time–frequency structure. Therefore, the study of active hearing as part of the echolocation system is possible by means of presenting artificially synthesized acoustic pulses to a dolphin so that these pulses simulate echo signals from different objects. Ignoring this important fact may cause misunderstanding of the terms "passive" and "active," because the main characteristics of active hearing are obtained from measurements in a seemingly passive mode [11].

The methodical aspects of experimental studies were described in our previous publications [1–3, 12], and, therefore, are not additionally discussed in this paper.

SENSITIVITY OF ACTIVE HEARING

The specificity of active hearing must manifest itself in a high sensitivity to individual short pulses and in an absence of time summation beyond the interval of 500 μs. Figure 2 shows the experimental dependence of the click hearing threshold (with respect to the threshold measured at a repetition rate of 3 PPS (pulses per second)) on the repetition rate F for two bottlenose dolphins [4].

For the dolphins being tested, the absolute sensitivity at F = 3 PPS was found to be $(4.4 \pm 0.4) \times 10^3$ μPa or (73 ± 0.8) dB and $(5.6 \pm 0.5) \times 10^3$ μPa or (75 ± 0.7) dB, respectively. At F = 30 PPS and a synthesized click duration of about 15 μs, the sensitivity values for these dolphins were $(3.0 \pm 0.3) \times 10^3$ μPa or (69.5 ± 0.9) dB. Note that these thresholds were close to those estimated from the evoked potentials of the auditory cortex of a bottlenose dolphin with the use of broadband stimuli (clicks) [13]. The minimal threshold value was found to be equal to 3×10^3 μPa (69.5 dB). The decrease in the hearing threshold because of the summation reached, on the average, 1.2 dB per doubling of the pulse repetition rate for F growing from 3 to 100 PPS. This summation was much lower than the energy one observed for humans (3 dB per doubling of the repetition rate) [14].

If the summation of pulses repeated F times per second occurs with a time constant τ, the signal will build up according to the law

$$I_F = I_1/\{1 - \exp(-1/F\tau)\}, \quad (1)$$

where I_F is the intensity of the signal at a repetition rate F and I_1 is the signal intensity at a certain initial repetition rate. For $F\tau \gg 1$, the intensity linearly grows with the repetition rate: $I_F = I_1 F\tau$; i.e., it increases by 3 dB when F is doubled. Assume that the signal is detected when its intensity exceeds a certain threshold I_0. Then, for the intensity I_F to be unchanged and equal to the threshold value I_0, it is necessary to compensate for the summation in the auditory system. For this purpose, I_0 should be taken in the form

$$I_F = I_0; \text{ or } I_1 = I_0\{1 - \exp(-1/\tau F)\}. \quad (2)$$

In Fig. 3, the dashed line represents curve *2* at τ = 200 μs (on a logarithmic scale). From the data shown in Fig. 2, it follows that, in the main operating range of repetition rates of echolocation clicks produced by bottlenose dolphins, which extends from 12 to 60 PPS, the summation of echo signals does not exceed 4 dB. Hence, presumably, it does not noticeably affect the increase in the noise immunity of active hearing against noise masker, compared to the effect of an increase in the echolocation click intensity. The latter increase,

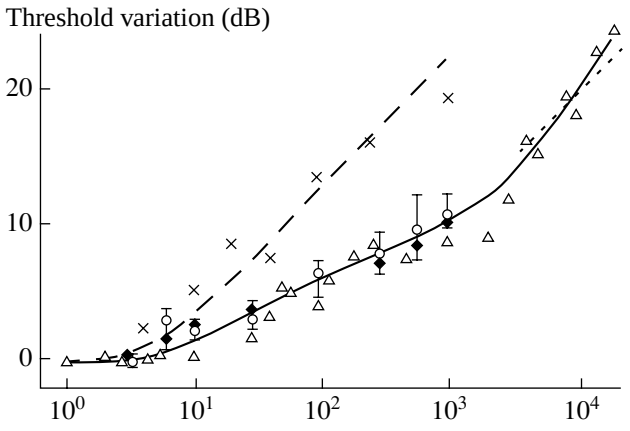


Fig. 2. Dependence of the variation in the hearing thresholds (in decibels) for synthesized pulses on the pulse repetition rate *F*: empty and full circles represent our data for two bottlenose dolphins, triangles represent the data taken from [16], and crosses refer to humans [14].

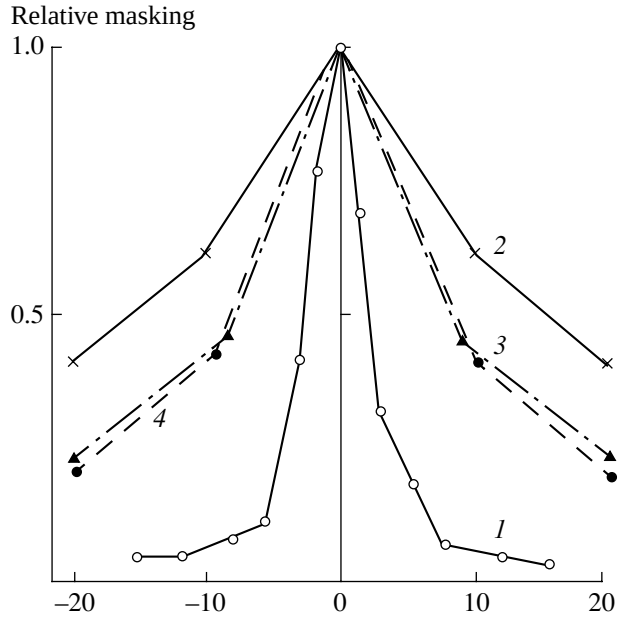


Fig. 3. Dependence of masking on the azimuthal separation of signal source S and noise source N. The abscissa axis represents the azimuth difference (in degrees) between signal S and noise N sources (see Fig. 2), and the ordinate axis, the effect of masking of the signal by noise with respect to that obtained with coincident S and N: (*1*) our data [27], (*2*) data for a tone signal of 80 kHz [28], (*3*) the same for a tone signal of 120 kHz, and (*4*) the directivity pattern for a tone burst with a carrier frequency of 80 kHz and a duration of 40 μs [29].

according to our data, may reach 52 dB and higher when the noise level grows [15]. Figure 2 also presents the data [16] obtained by the same method from one of the bottlenose dolphins under study but in a wider range of pulse repetition rates (from 1 to 20 000 PPS).

The comparison of these data with ours shows that, in the range of 3 to 1000 PPS, both groups of data agree well with each other. In addition, from the data presented in [16], one can see that, at $F = (3–4) \times 10^3$ pulse/s, the effect of summation drastically increases, and the summation becomes close to the energy one. The time constant of summation is close to 250 μs, which is shown in Fig. 2 by the dotted line. Au *et al.* [17, 18] obtained an integration time of 264 μs, which agrees well with our estimates. In their experiment, the detection threshold was measured for a pair of clicks in noise as a function of the time interval between them. When the interval was varied from 200 to 300 μs, the threshold increased by 3 dB, and a further increase in the click spacing did not cause any changes in the threshold value. This result testified to the absence of summation for click intervals exceeding 300 μs.

The summation time constants obtained in our experiments [1, 3] and in the experiments of Au, Moore, and Pawloski [17] were about 250 μs, while Johnson's experiments [19] on tone burst summation provided a noticeably different time constant of several tens of milliseconds. This discrepancy testifies to a difference in the auditory processing of short broadband clicks and narrow-band tone signals. Zaslavskii and Zanin [20] studied the time summation of acoustic clicks in a harbor porpoise. They used either pulses of standard shape with a duration of 40 μs or noise bursts in a frequency band of 80–180 kHz with a duration of 60 μs. As in bottlenose dolphins (Fig. 2), the curve of the time summation of the pulse sequence of a harbor porpoise consists of two parts: from 10 to 10^3 PPS and from 2×10^3 PPS and higher. Within the first part, the threshold decrease due to summation for standard-shaped pulses is 1.2–1.3 dB per doubled *F*. Within the second part, the threshold decrease reaches 5 and 4 dB per doubled *F* for noise pulses and standard-shaped pulses, respectively. It should be noted that the rate of threshold decrease was reliably determined to be higher than in the case of energy summation. At a pulse repetition rate of more than 3000 PPS, the rate of threshold decrease was also noticeably higher than 3 dB per doubled *F*. This means that the summation of the group of pulses falling within an interval of 250 μs is partially coherent.

Above, we considered the threshold summation of a sequence of pulses. Since, in the auditory system, the summation of stimuli usually decreases with the growth of their level above the hearing threshold (i.e., the time constant of summation decreases), one should expect an almost complete absence of superthreshold summation in dolphins in the range of pulse repetition rates from 1 to 1000 PPS. This means that the auditory detection of a target should mainly occur by a single echo pulse. The time summation measured by the brainstem evoked potentials [21] was observed for the durations up to 200–300 μs, which agrees well with the estimates

obtained from behavioral experiments for the integration time in the auditory system of bottlenose dolphins.

PULSE DISCRIMINATION BY INTENSITY AND TIME INTERVAL

In the measurement of differential intensity thresholds for short pulses, the specificity of active hearing should manifest itself in low threshold values. During the measurements, the signals to be compared were simultaneously emitted on both sides of a separating net thirty times a second. A higher-intensity pulse was randomly emitted to the left or to the right of the net. The task for the dolphin was to determine on which side this pulse was emitted. The differential intensity threshold (DIT) was defined as the intensity ratio of the stronger signal to the weaker one (in decibels), at which 75% correct responses were obtained. The measurements were performed for several signal levels above the hearing threshold. The results of the measurements are shown in Table 1.

Table 1. Average values of DIT and 90% confidence intervals for these values (in decibels)

	Signal level above the hearing threshold (dB)			
	5	20	36	45
Dolphin 1	2.5 ± 0.5	2.1 ± 0.8	2.1 ± 0.8	1.3 ± 0.6
Dolphin 2	1.7 ± 0.4	0.8 ± 0.1	0.7 ± 0.1	–

Table 2. Difference TPI and differential TPI for pulsed signals and 90% confidence intervals for their average values

Standard interval (μs)	50	100	200
Difference TPI (μs)	4.8 ± 0.8	9.4 ± 1.7	11.5 ± 2.4
Differential TPI (%)	9.6 ± 1.6	9.4 ± 1.7	5.7 ± 1.2
Differential TPI (dB)	0.4 ± 0.1	0.4 ± 0.1	0.2 ± 0.1

One can see that the DIT for clicks is close to 2 dB above the hearing threshold. As the signal level increases to 45 dB above the hearing threshold, the DIT decreases by a factor of 2–2.5 and reaches a value of about 1 dB. From the data obtained for bottlenose dolphins, it follows that, first, the DIT for clicks approximately coincides with the DIT for tone and noise signals, which testifies to a high intensity resolution of active hearing. Our data agree well with the estimates of the DIT by evoked potentials at the inferior colliculus [22] and auditory cortex [13]. Second, our DIT values reasonably agree with the estimates obtained from the data on auditory masking of one pure tone by another (1.0, 0.35, and 2.0 dB) [23]. Third, the measured DIT closely coincides with the relative target strength (~1 dB) of targets at the discrimination threshold, when the echo signal contains only primary echoes [2]. All these results confirm that the active hearing of dolphins is well adjusted to discriminating the intensities of pulsed stimuli.

Since bottlenose dolphins use clicks for echolocation, the echo signals returning from a variety of objects have the form of sequences of two or more clicks, which may differ in pulse separation. Therefore, the measurement of differential thresholds for the interval between pulses is of interest from the point of view of determining both the characteristics of active hearing and the discriminating abilities of the dolphin's echolocation system.

In the experiment on the discrimination between different pulse intervals, pairs of pulses were simultaneously emitted at a rate of 30 PPS on both sides of the separating net. The task of the dolphin was to choose the side where the pulses were produced with a greater interval between them. The standard intervals were 50, 100, 200, and 500 μs. The difference threshold in pulse interval (difference TPI) was determined as the difference between the intervals under comparison at 75% correct responses. The differential threshold in pulse interval (differential TPI) was determined as the ratio of the difference threshold to the standard interval (in percents or in decibels). At standard intervals of 50 and 100 μs, the differential TPI for one of the dolphins was found to be about 0.4 dB in both cases (Table 2).

At a standard interval of 200 μs, two of sixteen measurements of the difference TPI provided the values of 21.5 and 25.5 μs, which corresponded to a differential TPI of 10.7 and 12.7% (0.4 and 0.5 dB), respectively. At a standard interval of 500 μs, the dolphin was incapable of discriminating between the pulse intervals. Only once, an estimate of 34 μs was obtained for the difference threshold, which corresponds to a differential threshold of 6.8% (0.3 dB).

In our previous experiment [24], the discrimination thresholds for pulse intervals were measured with another bottlenose dolphin. The time intervals between pulses were taken within 50–100 μs at a step of 10 μs, and also longer intervals of 150, 180, and 200 μs were used. The average differential TPI were almost constant and varied from 5 to 8%, which is somewhat below the values presented in Table 2. The minimal difference thresholds reach the values of 1–2 μs. This agrees well with the data of echolocation experiments, in which it was demonstrated that, for the target discrimination, the dolphin had to discriminate between pulse intervals differing by only 1.8 μs [2]. Here we again observe the coincidence of the characteristics of active hearing in discriminating real echo signals (in the process of echolocation analysis) and in discriminating between artificial stimuli synthesized to simulate such echo signals.

At standard intervals of 180 μs and, especially, 200 μs, the discrimination was unstable. At the 180-μs

Table 3. Discrimination between pulse intervals exceeding 300 μs

Intervals compared (μs)	300	500	500	500	500	500	1000	1000	1000	1800	2000
	400	500	556	600	700	1200	1100	1200	1500	2000	2000
Percentage of correct discriminations	56	50	53	64	46	60	36	50	45	57	43

interval, the threshold could be determined in four cases out of six, and in two of these four cases, the thresholds were obtained not in the usual way, i.e., by gradually decreasing the greater test interval to the standard (smaller) interval, but in the opposite way: the standard (greater) interval was fixed and the smaller (test) interval was gradually increased toward the standard value. Despite the difficulties in threshold measurement, the threshold value at a standard interval of 180 μs was still fairly small. At a standard interval of 200 μs, the threshold was determined in three cases out of six: once "from above" and twice "from below," and the resulting values of the difference TPI were rather high (25–32 μs).

Table 3 shows the data characterizing the ability of one of the bottlenose dolphins to discriminate between pulse intervals exceeding 300 μs. From Table 3, it follows that such intervals are not discriminated (the percentage of correct discrimination is close to 50% in different experiments). Similar results were obtained by us with another bottlenose dolphin [24]. We also tried to observe the discrimination between pairs of pulses of a rectangular shape with a duration of 50 μs and a carrier in the form of white noise in a frequency band of 50 kHz. The repetition rate was 2 PPS [10]. The intervals between pulses in the pairs suggested for comparison were either 100 and 1100 μs or 2000 and 250 μs. The dolphin could not be taught to discriminate between these intervals.

On the basis of these results, one can conclude that a sequence of echo pulses is perceived by a dolphin as a single auditory image if it falls within a certain time interval not exceeding 200–300 μs. This critical time, T_{cr}, determines the maximal size (20 cm) of the object whose echo signal is perceived by the dolphin as a merged auditory image. This time coincides with the duration of the interval within which an efficient threshold summation of repeated pulses takes place.

In the introduction, it was mentioned that the necessity to perform a comprehensive analysis of echo signals from ecologically significant objects (fish) requires especially intricate mechanisms of signal processing within time intervals smaller than 500 μs. Here, we observe a manifestation of the operation of these mechanisms by the example of the discrimination between pulse intervals in pairs of pulses.

NOISE IMMUNITY OF ACTIVE HEARING

The effect of masking one acoustic signal by another is widely used in hearing studies on both humans and animals to reveal the interaction between two stimuli, signal and noise. In our experiments, we used the effect of masking to determine the characteristics of active hearing in the situations where one pulse was masked by another and also where a true echo signal was masked by a noise pulse synchronized with the dolphin's own echolocation click.

In the experiment on noise immunity of active hearing, two sequences of pulses produced on both sides of a separating net were simultaneously presented to the dolphin under investigation. One source generated a sequence of single pulses thirty times a second. Another source generated a sequence of pairs of pulses in such a way that the first pulse of a pair coincided in time with a single pulse synchronously emitted on the other side of the net. The amplitude and delay of the second pulse in a pair was varied in the course of the experiment. The emission of a single pulse and a pair of pulses occurred at random on both sides of the net. The task of the dolphin was to detect the presence of the second pulse in a pair. The time interval T between paired pulses was varied in different series of measurements from 0 to 10 ms. The measurement error for T did not exceed 3 μs, and the measurement error for the amplitude ratio was about 2 dB.

Experiments with one of the dolphins showed [1–5] that, for $T > 500$ μs, the masking thresholds were close to the absolute threshold for a single pulse within the experimental error. This means that auditory excitations do not interact when they are separated in time by approximately 500 μs. By contrast, for $T < 300$ μs, a strong interaction of stimuli is observed, which, in the case under consideration, manifests itself as a strong masking.

Our data are in good agreement with the results of studying the direct temporal masking of evoked potentials in the brainstem of dolphins, when noise arrives before the signal [8, 21]. This corresponds to the echolocation of an object at a small distance from the dolphin with noise being represented by the own echolocation click of the dolphin. In this sense, our experiments on forward masking correspond to the case of a weak echo masked by a stronger one delayed with respect to it.

Evoked potentials were completely separated when the delay between noise and signal was 2–5 ms. When

the dolphin's own echolocation click served as masker, a time separation of evoked potentials caused by the two pulses was observed for even shorted delays of about 1 ms. The reconstruction of the first peak of the evoked potential occurred in 100% of the cases with a delay of only 500 μs [21]. The minimal response to the signal at equal signal and masker amplitudes was observed at a pulse interval of 200 μs.

Similar threshold values (200–300 μs) of pulse intervals, at which the response to the second click of a pair could be reliably detected, were obtained by us in electrophysiological experiments on the harbor porpoise [25]. Thus, in the experiment with a backward time masking, we again arrive at the fact that the mechanisms of auditory processing of echo pulses are different for the case when the pulses forming the stimulus fall within a time interval of about 300 μs and for the case when they fall outside this interval.

The specificity of active hearing of bottlenose dolphins should presumably manifest itself most strongly in the perception of echo signals from objects. Observations of the echolocation abilities of dolphins under strong reverberation conditions (detection or identification of objects near the wall or the bottom) show that dolphins possess a high noise immunity with respect to reverberation. These observations gave rise to the assumption that Odontocetes possess special mechanisms providing detuning from reverberation pulses or from their own echolocation clicks [2]. For example, this may be a time gating mechanism, whose operation includes blocking of the auditory system at the instants of arrival of intense reverberation pulses and the enhancement of hearing sensitivity at the instants of arrival of weak echo signals.

To determine the specific features of active hearing in the presence of dolphin's own radiation, we carried out experiments in which dolphins detected a steel ball 40 mm in diameter against a background of artificial reverberation pulses [1–5]. The latter were produced by emitting pulses that simulated the echolocation clicks of a dolphin in shape and spectrum. The time relation of the interfering pulses to the echolocation clicks of the dolphin and the echo signal from the ball was achieved by means of the initiation of the interfering pulses by the echolocation clicks of the dolphin. Controlling the delay of the interfering pulse emission, it was possible to vary the seeming position of the source of interference with respect to the source of echo (the ball). The echo-to-noise ratio in this experiment was -50.0 ± 3.5 dB. The results of the experiment showed that, for $T < -300$ μs and $T > 500$ μs, the detection of the ball was quite reliable. This means that the interfering pulse did not mask the echo signal arriving from the ball. The time resolution determined in this experiment almost coincides with that obtained in the absence of echolocation clicks (a backward masking). This suggests that the high noise immunity of active hearing with respect to reverberation is achieved in dolphins through the high time resolution of active hearing rather than through the aforementioned hypothetical strobing mechanism.

Moore *et al.* [26] determined the value of the critical interval in an experiment on backward time masking under the echolocation conditions where noise was initiated by the echolocation click of the dolphin. The results of this experiment confirmed our data: the critical interval was estimated to be equal to 265 μs. In this experiment, we again encounter the critical interval in the natural mode of operation of dolphins' echolocation system.

Until this point, we considered masking with a time delay between signal and noise. In another experiment, for a pulse simulating an echolocation signal, we studied its masking by a spatially separated noise pulse [27]. For this purpose, two pairs, S and N, of spherical piezoceramic transducers were positioned on both sides of the net. Paired transducers N were placed at a constant distance of 0.3 m from the separating net and used for producing noise pulses. Paired transducers S were used for randomly emitting pulsed signals at the left or at the right of the net. These transducers were positioned symmetrically with respect to the net, but their distance from the net could be varied. The sound pressure level of noise pulses reached 114 dB relative to 1 μPa at the net end nearest to the dolphin. The noise pulses were emitted simultaneously on both sides of the net (synchronously with the clicks produced by one of the transducers of pair S). The repetition rate of all pulses was 30 PPS. The pulse duration was close to 10 μs, and the spectrum maximum corresponded to approximately 80 kHz. The initial position of the dolphin was at the beginning of the net. The dolphin detected the signal and moved to the side of the net where it thought the signal source was located. At a given distance of the transducer from the net, the signal detection threshold was determined by a level of 75% correct responses.

The decrease in masking because of the azimuthal separation of transducers S and N is characterized by Fig. 3 (curve *1*). The abscissa axis represents the azimuth difference between signal source S and noise source N in degrees, and the ordinate axis represents the ratio $20\log(P_\theta/P_0)$, where P_θ and P_0 are the peak sound pressures of the signal at some azimuth θ and at $\theta = 0°$, respectively. The results demonstrate a high noise immunity of active hearing in the case of the spatially separated signal and noise sources. When the signal and noise sources are at the same point ($\theta = 0°$), the effect of masking is 43.6 dB. When the signal source is spatially separated from the noise source by as little as 3.2°, the effect of masking is reduced by 8–10 dB, and when the separation reaches 15°, it decreases on the average by 30 dB. The specificity of active hearing manifests itself in that the noise immunity proves to be higher for clicks simulating echolocation ones than for high-frequency tone bursts (Fig. 3). At 0.4 of maximum

value, the widths of curves *1*, *2*, and *3* are 6°, 23°, and 40°, respectively, although the frequency corresponding to the spectrum maximum of the pulsed signal (80 kHz) coincides with the carrier frequency of one of the signals.

Curves *1–3* characterize the spatial resolution of hearing, and, in this respect, they can be compared with hearing directivity patterns, which describe the anisotropy of hearing sensitivity. As was shown in [28, 29], the directional characteristic of reception narrows when the signal of the same frequency becomes shorter (curve *4* in Fig. 3).

Thus, in this experiment, the functional specificity of active hearing manifests itself in that the pulse simulating the echolocation click of the dolphin proves to be the most effective stimulus: it allows the active hearing system to perform the most efficient spatial selection of signal against noise when the space–time structure of masking noise coincides with that of the signal.

CRITICAL INTERVAL OF ACTIVE HEARING

The specific features of auditory analysis of acoustic pulses, which manifest themselves in the considerable difference between the mechanisms of auditory processing of paired pulses separated by more than 500 μs and less than 200 μs, made it necessary to measure the critical interval T_{cr} with a higher accuracy [1, 2, 30].

In the measurements of the duration of the critical interval T_{cr}, the task set before the bottlenose dolphin was to discriminate between the standard pair of pulses separated by an interval of 100 μs and the test pair of pulses separated by an interval varying from 150 to 400 μs. The sound pressure level of each of the pulses was 40 dB above the hearing threshold. The standard and test intervals were randomly produced to the left or to the right of a 7-m-long separating net. In this experiment, we used the method of a "special" presentation of a pair of pulses separated by a test interval on the background of the presentation of two "background" pairs of pulses with intervals of 100 and 150 μs. The animal should choose the side of the net where the pair of pulses with the greater interval was produced.

The results of the experiment are shown in Fig. 4. The abscissa axis represents the test interval in microseconds, and the ordinate axis, the percentage of correct determinations of the side (with respect to the net) where the transducer producing the pair of pulses with the test interval was positioned. The vertical segments at the experimental points show the 90% confidence intervals for the average values. The absence of such segments at some of the experimental points means that the corresponding confidence interval is equal to zero. As can be seen from Fig. 4, the percentage of correct responses given by the dolphin varies from 100% for $150\ \mu s < T_{test} < 200\ \mu s$ to zero for $T_{test} > 325\ \mu s$. The zero percentage means that, when the test interval was increased beyond the critical interval, the dolphin's choice of the positive stimulus was the pair of pulses separated by the standard (smaller) interval rather than the pair of pulses separated by the test (greater) interval. In our opinion, this occurred because an increase in the test interval beyond the critical interval lead to a decay of the merged auditory image of the test pair into two independent auditory images corresponding to individual pulses of this pair. Then, the pair of pulses with a standard interval of 100 μs proved to be a "more positive" stimulus than each single pulse of the test pair. A total loss of discrimination between pairs of pulses occurred at $T_{test} = 260 \pm 25$ μs, and the threshold discrimination (75% of correct responses) was observed at $T_{test} = T_{cr} = 230 \pm 40$ μs.

Thus, this experiment determines the value of the critical interval, within which the auditory images of two individual pulses are merged into a single auditory image of a pair. The main result of the experiment described above is not only the accurate measurement of the value of T_{cr} but also the determination of the fundamental meaning of the term "critical interval" as a certain characteristic time within which the interaction of acoustic pulses takes place and two auditory events are perceived as merged into one single image.

Evidently, the change from a single image of a pair to completely independent images of individual pulses occurs gradually but not necessarily in a monotonic way. It depends on a number of physical and physiological factors, such as the sound pressure level of the pulses above the threshold, the amplitude ratio of the pulses, the subjective features used by the dolphin in making decisions, etc. For example, one can expect that, at $T = T_{min}$ (the exact value of T_{min} is to be determined later), the auditory image of a pair of pulses will have the form of a single image, and in the range of intervals $T_{min} < T < 1000$ μs, this image will gradually decay. In other words, the beginning of the decay of a single image can already be detected at $T = T_{min}$, but, on the other hand, even at $T = 500$ μs, traces of a single image of two pulses may still be observed. Hence, in specific experiments, one or another value of T_{cr} can be obtained, depending on whether the dolphin discriminates between pairs of pulses or more complex sequences of echo pulses and whether it uses the features characterizing a single image or its individual components.

The critical interval property also means that the mechanisms of the auditory analysis of echo pulses falling within this interval prove to be different from the mechanisms of auditory analysis of pulses falling outside this interval.

In this connection, one should note the following facts.

First, a considerable increase in the efficiency of the threshold summation occurred when the interval between neighboring pulses became smaller than 300 μs (Fig. 2). The estimate of T_{cr} by the change in the

slope of the summation curve yields a value lying within 300 to 500 μs.

Second, without special training, the dolphins under study could discriminate between the paired pulses only within the limits of the critical interval. Estimates of T_{cr} from these data yield values of T_{cr} within 200–500 μs. Later, we managed to teach one of the dolphins to distinguish between pairs with pulse intervals exceeding the critical interval, but it was evident that, in this case, the criterion used by the dolphin was the difference between the pulse intervals in the pairs suggested for comparison. The change from discrimination of pairs with intervals greater than critical to discrimination of pairs with intervals smaller than critical was always accompanied by a change in the criterion of decision and an increase in the scatter of data. All this is typical of a boundary corresponding to a change in the criterion (feature) used for discrimination [31].

Third, both forward and backward time masking of one pulse by another occurs within the limits of the critical interval. The estimates of the boundaries of T_{cr} from the masking curves yield the following result: 100 μs < T_{cr} < 1000 μs.

Fourth, the direct and inverse masking of the echo signal by a pulse synchronized with the echolocation click were observed for two bottlenose dolphins only within the critical interval. The estimates of T_{cr} obtained from this experiment yield 200 μs < T_{cr} < 500 μs.

Fifth, a direct measurement of the width of the critical interval as an interval where auditory images of individual pulses merge into a single image of the pair provided the estimate of $T_{cr} = 230 \pm 40$ μs.

In the behavioral experiments with bottlenose dolphins, by the method of selective adaptation, we tested the hypothesis proposed by M.N. Sukhoruchenko about the presence of specialized channels for the time interval selection in the auditory system of dolphins [32, 33]. We obtained the dependences of the detection threshold for the second pulse of the test pair on the pulse interval in pairs used as adapting ones. It was shown that, for test pairs with pulse intervals up to 200 μs, the curves exhibit a peak of adaptation when the test interval coincides with the adapting interval. For a test interval of 500 μs, no dependence of the aforementioned threshold on the adapting interval was observed.

ON THE STRUCTURE AND MECHANISMS OF DATA PROCESSING WITHIN THE CRITICAL INTERVAL

Thus, the fundamental property of the mechanism of data processing within T_{cr} is the interaction of the echo highlights and the formation of a merged auditory image from them. Specific manifestations of this property may be different, for example, it may be an increase in the summation efficiency or an increase in the sensitivity to the spectral and time structures of stimuli. The limits of manifestation of this property also depend on which subjective quality of the combined auditory image is used by the dolphin in a given experiment.

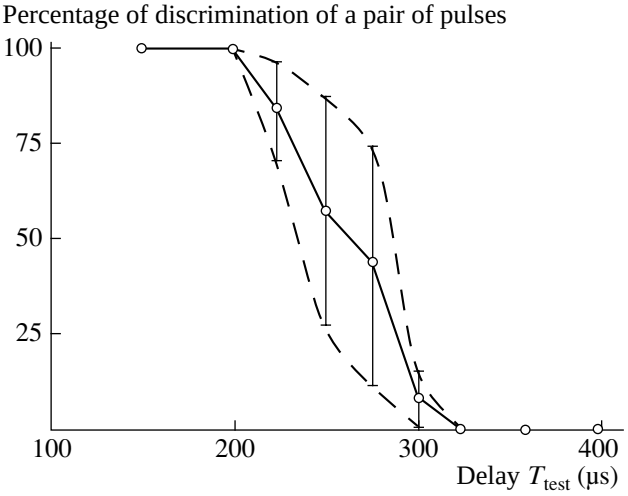


Fig. 4. Percentage of correct discriminations for the pair of pulses separated by the interval T_{test} and a pair of pulses separated by an interval of 100 μs versus the value of the interval T_{test}. The vertical segments indicate the 90% confidence intervals for the average values [30].

In this respect, the mechanism of data processing for auditory events within T_{cr} is similar to the mechanism of data processing within the critical band Δf_{cr}: a confluence of auditory images corresponding to individual spectral components, which fall within the critical band, into a single auditory image. Presumably, this similarity is not purely external. One should expect that the mechanisms of auditory analysis within T_{cr} are as complicated as the mechanisms known for critical bands. Within T_{cr}, a complete set of embedded mechanisms may exist, each mechanism corresponding to its own subinterval smaller than T_{cr}. As will be seen below, the experimental facts obtained by us testify to the existence of such subintervals.

Let us first discuss the results of the experiments performed to reveal the mechanisms of auditory signal processing within T_{cr} [34]. The first question we tried to answer was whether the dolphin analyzes the time profile of the echo signal within the critical interval, i.e., whether it uses all data available, or it performs the discrimination on the basis of some reduced description of the echo signal, e.g., its energy spectrum. In the first experiment, the dolphin was offered two pairs of pulses for discrimination: $S_1(t) = aS(t) + bS(t - T)$ and $S_2(t) = bS(t) + aS(t - T)$, where $S(t)$ is the time profile of a single pulse, a and b are positive numbers ($a > b$), and T is the time delay between the pulses. Pairs $S_1(t)$ and $S_2(t)$ have identical energy spectra. For the dolphin, the lack or the presence of the ability to discriminate between pairs $S_1(t)$ and $S_2(t)$ means that its auditory analysis is

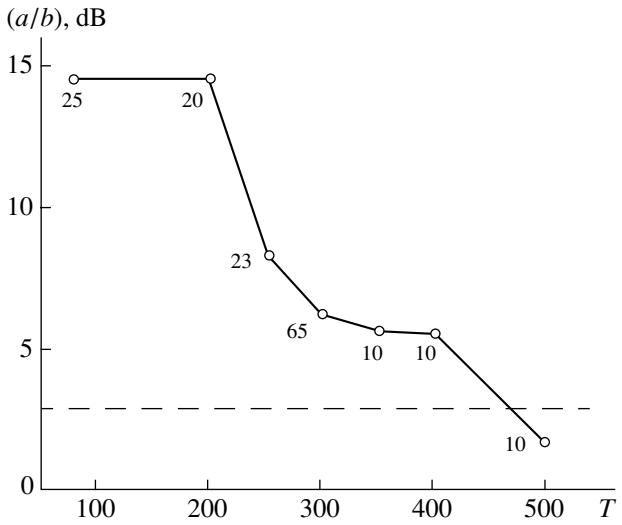


Fig. 5. Results of one of the experiments on the discrimination between pairs of pulses of forms S_1 and S_2. The horizontal axis represents the interval T in microseconds, and the vertical axis, the threshold ratio of pulse amplitudes in a pair, $20\log(a/b)$, in decibels. The signal level above the hearing threshold is 40–45 dB. The numbers near the experimental points indicate the numbers of tests used to obtain the corresponding threshold estimates. The dashed line shows the differential intensity threshold measured in the given experiment [34].

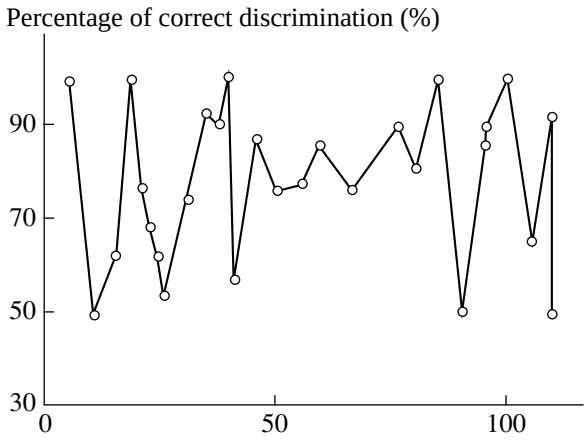


Fig. 6. Discrimination between pairs of pulses of forms S_1 and S_2. The horizontal axis represents the interval T in microseconds, and the vertical axis, the percentage of correct discriminations [34].

based on the use of only the energy spectrum of the pair of pulses (a reduced description) or on the use of the full description of the signal, respectively.

The results of the first experiment with such stimuli are presented in Fig. 5. The abscissa axis represents the interval T between pulses in microseconds, and the ordinate axis, the ratio $20\log(a/b)$ in decibels. Circles indicate the threshold values of $20\log(a/b)$ corresponding to 75% correct discrimination between stimuli. An exception is the point corresponding to 100% at T = 500 μs. Thus, part of the first quadrant between coordinate axes, which lies below the experimental curve, corresponds to nondiscrimination. As seen from Fig. 5, within T_{cr} the discrimination occurs when the amplitude ratio exceeds 14.5–15 dB. When T increases up to 250 μs, this ratio rapidly decreases reaching a value of 1.6 dB at T = 500 μs. We believe that, when $T > T_{cr}$, i.e., when two pulses forming one pair do not merge into a single auditory image, the dolphin discriminates between the pairs of pulses according to the amplitude difference between the first and second pulses in these pairs. By contrast, in the case of $T < T_{cr}$, the lack of discrimination for amplitude ratios below 14 dB testifies that, in this experiment, the dolphin cannot discriminate between stimuli with identical energy spectra.

This result may mean either the fundamental impossibility for the dolphin to discriminate between signals with identical energy spectra, i.e., a purely spectral mechanism of signal processing within T_{cr}, or the situation where the dolphin could not "tune in" to the corresponding time features. Therefore, the experiment on discrimination between pairs $S_1(t)$ and $S_2(t)$ was continued at a constant amplitude ratio of 10 dB with a variable interval T. Its results are presented in Fig. 6. At the very beginning of the experiment, a total absence of discrimination between pairs of pulses was observed at T = 10, 20, 40, and 200 μs (the corresponding numbers of tests were 30, 20, 20, and 60). However, in the subsequent tests, the dolphin could discriminate between pairs of pulses with intervals T equal to 20 (88%, 36 tests), 30, and 60 μs. The observation of a confident discrimination at these interval values suggests that the dolphin was able to find the features depending on the time structure of the stimuli.

Note that the dependence of the percentage of correct discriminations of pairs S_1 and S_2 on the interval T exhibits more or less pronounced oscillations with a period of T = 15 μs. As the experimentation with these stimuli was continued, the percentage of discrimination between S_1 and S_2 increased to 100%. This experiment uniquely testifies to the possibility of a temporal processing of stimuli within T_{cr}. Hence, within T_{cr}, both purely spectral processing and temporal processing of signals are possible.

It is known that the description used by the auditory system of humans noticeably differs from the "physical" description in the form of an energy spectrum of temporal profiles of the signal. As a subjective analog of the physical spectral description, one may consider the timbre-based description, i.e., the distribution of the signal loudness over the critical bands of hearing. Assuming that a bottlenose dolphin can use such a timbre-based description of the signal, we consider two possible mechanisms of spectral (timbre-based) discrimination between pairs of pulses with a pulse spacing $T < T_{cr}$ [34].

One of the mechanisms can be called differential, because, for making a decision about the interval discrimination, it is sufficient to analyze the difference in the signal loudness at the output of only one critical band. In the case of the threshold discrimination of intervals, when the differential threshold is $\Delta T/T < 1$, we arrive at the following expression for this mechanism: $\Delta T/T = \Delta f_{cr}/f^{\wedge}$, where Δf_{cr} is the width of the critical band of hearing at a certain frequency $f^{\wedge}$. This width is determined by the effective width Δf_{eff} of the spectrum of individual pulses forming the pair, the position of the spectral maximum of an individual pulse on the frequency axis, and the audiogram of the dolphin [34].

Thus, the spectral discrimination mechanism considered above leads us to the condition of constancy of the differential threshold within the time interval. However, this mechanism cannot function in the case of limiting pulse intervals, for which the distance between neighboring maxima of the energy spectrum of a pair of pulses is less than twice the critical band width. In this case, the spectral maxima of both pairs of pulses prove to be within the same critical band and cannot be resolved by the auditory system of the dolphin.

However, for the dolphin, it is also possible to use the difference in the timbres of two pairs of pulses, i.e., the difference in the loudness distribution over different critical bands lying within the effective spectral width of an individual pulse, Δf_{eff} [34]. If this spectral mechanism (we call it the integral one) is used by the dolphin, the differential threshold in time interval decreases in proportion to T while the difference threshold remains constant.

Let us consider the dependences of the differential thresholds in time interval on the interval duration for the differential and integral hypothetical mechanisms of auditory spectral processing and compare them with the experimental data. In Fig. 7, the solid line shows the dependence $\Delta T/T = \Delta f_{cr}/f^{\wedge}$. The dashed line shows the dependence $\Delta T/T = 1/(f_{up}T)$ at $\Delta f_{eff} = 50$ kHz (here, f_{up} is the upper limiting frequency of the spectrum of an individual pulse). Empty and full circles represent our previous experimental data, and crosses represent the results reported in [16] for the same animal. For $T <$ 100 μs, the experimental values of $\Delta T/T$ remain practically constant. For $T > 100$ μs, a rapid decrease in the differential threshold is observed. Thus, we conclude that the experimental data on the discrimination of pairs of pulses with different time intervals agree well with two hypothetical mechanisms of auditory spectral (timbre-based) processing within the critical interval.

The determination of the features used in the discrimination of pairs of pulses that may differ in both the pulse spacing (oscillations, or microstructure of the energy spectrum) and the individual pulse shape (envelope, or macrostructure of the energy spectrum) was performed in [35]. It was shown that dolphins use the microstructure of the spectrum to discriminate pairs of

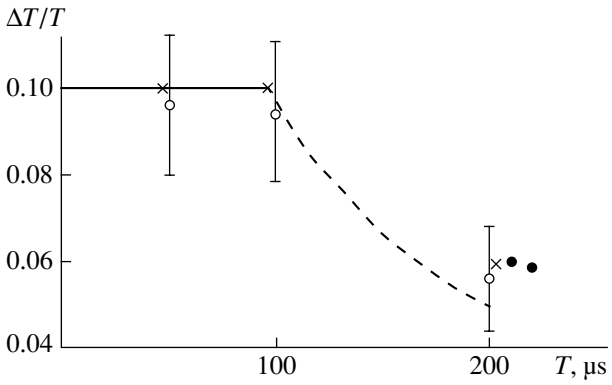


Fig. 7. Dependence of the differential threshold in the time interval $\Delta T/T$ on the duration of this interval [24, 30]. Crosses show the results obtained by A.V. Zanin and G.L. Zaslavskii for the same dolphin.

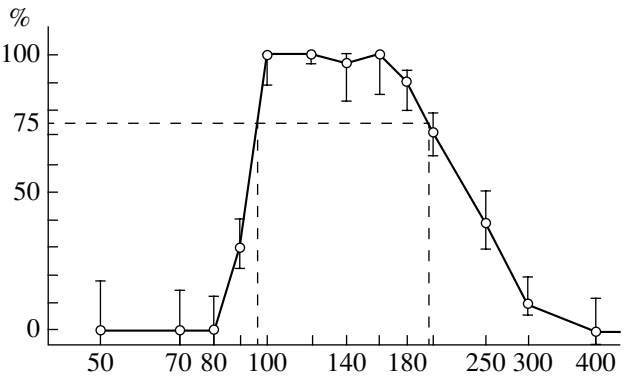


Fig. 8. Percentage of the responses of the bottlenose dolphin to test signals versus the pulse interval in microseconds [35]. The vertical segments indicate the confidence intervals for the 5% significance level. The vertical dashed lines indicate the boundary corresponding to the change from the macrostructure feature to the microstructure feature (at 97 μs) and the upper limit for the microstructure feature (the boundary of the critical interval in the given experiment).

pulses with time intervals from 97 to 200 μs, while for time intervals smaller than 97 μs, they use the macrostructure of the spectrum (Fig. 8). It was found that the macrostructure feature is dominant with respect to the microstructure feature. This means that the dolphin first tries to discriminate the pairs of pulses by the macrostructure of their spectra. If the difference in this feature is insufficient, the dolphin uses the microstructure feature of the spectra. The results of these thorough experimental studies show that, for discrimination, a dolphin may use two different features (processing mechanisms) in different subintervals within the critical interval.

An important argument in favor of the hypothesis that dolphins possess passive and active hearing may be the audiograms obtained for a harbor porpoise and an inia with the use of evoked potentials recorded on the surfaces of animal's heads [36]. The audiograms have

two pronounced minima. The first minimum occurs at low frequencies of 25–30 kHz for both species. We believe that this frequency band is used by passive hearing of these dolphins. The second minimum occurs at high frequencies of about 110 kHz for harbor porpoise and 70–80 kHz for inia. We believe that this frequency band is used by active hearing of inia and harbor porpoise. The separation of frequency bands used by passive and active hearing manifested itself most clearly in this case, because the echolocation clicks of these species noticeably differ from those of bottlenose dolphins. They are characterized by higher frequencies, narrower frequency bands, and greater durations: 200–250 μs in the 10-kHz frequency band from 95 to 105 kHz in inia [37] and 130–260 μs in the 20-kHz frequency band from 120 to 140 kHz in harbor porpoise [38–40].

CONCLUSIONS

A generalization of the data presented above suggests the following conclusions. The auditory system of dolphins is well adjusted to analyzing acoustic stimuli that consist of a single click or a sequence of clicks simulating echolocation pulses or echoes from underwater objects. This adjustment manifests itself in (i) a high sensitivity of hearing to pulsed signals; (ii) a weak threshold summation of pulses in the main operating range of repetition rates of echolocation clicks (a threshold decrease of 1.2 dB with doubling of repetition rate), which testifies to a single-pulse character of auditory detection, i.e., detection by a single echo signal; (iii) a high differential sensitivity in the short pulse discrimination by intensity (~1 dB) (the values of differential intensity thresholds for short pulses coincide with those for long noise and tone bursts and echoes); and (iv) a noticeable difference in the mechanisms of auditory processing used for the echo pulse components lying within and outside the critical interval.

Thus, our hypothesis that dolphins possess two functionally (and, possibly, anatomically to some extent) separated systems of auditory perception, namely, active and passive hearing, is confirmed by the whole set of available experimental data.

REFERENCES

1. V. A. Vel'min and N. A. Dubrovskiĭ, Dokl. Akad. Nauk SSSR **225**, 470 (1975).
2. V. M. Bel'kovich and N. A. Dubrovskiĭ, *Sensory Basis of Cetacean Orientation* (Nauka, Leningrad, 1976).
3. V. A. Vel'min and N. A. Dubrovskiĭ, in *Marine Mammals: Results and Methods of Investigation* (Nauka, Moscow, 1978), pp. 90–98.
4. N. A. Dubrovsky, in *Sensory Abilities of Cetaceans*, Ed. by J. Thomas and R. Kastelein (Plenum, New York, 1990), pp. 233–254.
5. N. A. Dubrovsky, in *Bottlenose Dolphin of the Black Sea*, Ed. by V. E. Sokolov (Nauka, Moscow, 1997), pp. 544–574.
6. G. Bekesy, *Experiments in Hearing* (McGraw-Hill, New York, 1960).
7. D. N. Jacobs, J. Acoust. Soc. Am. **52**, 696 (1972).
8. T. N. Bullock and S. H. Ridgway, J. Neurobiol. **3** (7), 79 (1972).
9. V. I. Markov, in *Proceedings of IX All-Union Acoustical Conference* (Moscow, 1977), p. 41.
10. É. Sh. Aĭrapet'yants and A. I. Konstantinov, *Echolocation in Nature* (Nauka, Leningrad, 1974).
11. W. W. L. Au, *The Sonar of Dolphins* (Springer, New York, 1993).
12. V. K. Bezrukov, R. É. Gassko, N. A. Dubrovskiĭ, and G. L. Zaslavskiĭ, in *Proceedings of X All-Union Acoustical Conference* (Moscow, 1983), p. 145.
13. A. Ya. Supin, L. M. Mukhametov, E. F. Ladygina, *et al.*, *Electrophysiological Study of Dolphin's Brain* (Nauka, Moscow, 1978).
14. J. Zwislocki, in *Handbook of Mathematical Psychology*, Ed. by R. D. Luce, R. R. Bush, and E. Galanter (Wiley, New York, 1963), Chap. 17.
15. V. P. Babkin and N. A. Dubrovsky, Tr. Akust. Inst. Akad. Nauk SSSR, No. 7, 29 (1971).
16. A. V. Zanin, G. L. Zaslavskiĭ, and A. A. Titov, in *Proceedings of IX All-Union Acoustical Conference* (Moscow, 1977), p. 21.
17. W. W. L. Au, P. W. B. Moore, and D. A. Pawloski, J. Acoust. Soc. Am., Suppl. **80**, S107 (1986).
18. W. W. L. Au, P. W. B. Moore, and D. A. Pawloski, J. Acoust. Soc. Am. **83**, 662 (1988).
19. C. S. Johnson, J. Acoust. Soc. Am. **43**, 757 (1968).
20. G. L. Zaslavskiĭ and A. V. Zanin, in *Marine Mammals* (Moscow, 1978), pp. 128–129.
21. V. V. Popov and A. Ya. Supin, in *Electrophysiology of the Sensory Systems of Marine Mammals* (Nauka, Moscow, 1986), pp. 85–106.
22. T. H. Bullock, A. D. Grinnell, E. Ikezono, *et al.*, Z. Vgl. Physiol., 117 (1968).
23. C. S. Johnson, J. Acoust. Soc. Am. **49**, 1317 (1970).
24. N. A. Dubrovsky, L. K. Rimskaya-Korsakova, and M. N. Sukhoruchenko, in *Abstracts of Papers of VII All-Union Workshop on Marine Mammals* (Simferopol, 1978), p. 113.
25. N. G. Bibikov, L. K. Rimskaya-Korsakova, A. V. Zanin, and N. A. Dubrovsky, in *Electrophysiology of the Sensory Systems of Marine Mammals* (Nauka, Moscow, 1986), pp. 56–84.
26. P. W. B. Moore, R. W. Hall, W. A. Friedl, and P. E. Nachtigall, J. Acoust. Soc. Am. **76**, 314 (1984).
27. V. A. Vel'min and N. A. Dubrovskiĭ, in *Proceedings of IX All-Union Acoustical Conference* (Moscow, 1977), p. 5.
28. K. A. Zaĭtseva, A. I. Akopian, and V. P. Morozov, Biofizika **20** (3), 519 (1975).
29. A. I. Akopian, K. A. Zaĭtseva, V. P. Morozov, and A. A. Titov, in *Proceedings of IX All-Union Acoustical Conference* (Moscow, 1977), p. 9.
30. V. A. Vel'min and N. A. Dubrovsky, Akust. Zh. **22**, 351 (1976) [Sov. Phys. Acoust. **22**, 351 (1976)].

31. N. A. Dubrovsky, L. K. Rimskaya-Korsakova, and M. N. Sukhoruchenko, in *Marine Mammals* (Moscow, 1978), pp. 114–115.
32. N. A. Dubrovsky and M. N. Sukhoruchenko, in *Abstracts of Papers of IX All-Union Workshop on Study, Protection, and Rational Usage of Marine Mammals* (Arkhangelsk, 1986), p. 127.
33. M. N. Sukhoruchenko and N. A. Dubrovsky, in *Sensory Systems and Behavior of Aquatic Mammals* (Moscow, 1991), pp. 95–97.
34. N. A. Dubrovsky, P. S. Krasnov, and A. A. Titov, in *Proceedings of 9th All-Union Acoustical Conference* (Moscow, 1977), p. 25.
35. N. A. Dubrovsky, T. V. Zorikov, O. Sh. Kvizhinadze, and M. M. Kuratishvili, Akust. Zh. **37**, 18 (1991) [Sov. Phys. Acoust. **37**, 485 (1991)].
36. V. V. Popov and A. Ya. Supin, in *Sensory Abilities of Cetaceans*, Ed. by J. Thomas and R. Kastelein (Plenum, New York, 1990), pp. 405–416.
37. C. Kamminga, F. J. Engelsma, and R. P. Terry, in *Proceedings of 8th Biennial Conference on Biology of Marine Mammals* (Pacific Grove, California, 1989).
38. B. Möhl and S. Andersen, J. Acoust. Soc. Am. **54**, 1368 (1973).
39. C. Kamminga and H. Wiersma, Aquat. Mammals **8**, 41 (1981).
40. Y. Hatakeyama and H. Soeda, in *Sensory Abilities of Cetaceans*, Ed. by J. Thomas and R. Kastelein (Plenum, New York, 1990), pp. 269–281.

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 318–322. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 384–388.

BIOLOGICAL ACOUSTICS

The Use of Wideband Signals in the Acoustic Monitoring of Hydrobionts

V. A. Zhuravlev*, A. A. Mazanikov, and A. N. Neronov**

** Andreev Acoustics Institute, Northern Branch, Russian Academy of Sciences, ul. Pochtovaya 3, Severomorsk, Murmansk oblast, 184600 Russia*

e-mail: sfakin@severm.mels.ru

*** Murmansk State Technical University, ul. Sportivnaya 13, Murmansk, 183010 Russia*

e-mail: Alexander.Mazanikov@mstu.edu.ru

Received November 18, 2003

Abstract—A method is presented for the coherent detection of the signal scattered by a hydrobiont moving across the path of sound propagation from a wideband source against the background of the sum of an intense insonification signal and additive noise. The method consists of the calculation of a cross-correlation function of the signal at the receiver input and the scattered signal model constructed from the received signal by introducing the square-law time delay. It is shown theoretically and with numerical modeling that such a processing ensures a gain in the signal-to-noise ratio that is proportional to the square root of the bandwidth multiplied by the observation time. The efficiency of the method is demonstrated by the results of an experiment. © *2004 MAIK "Nauka/Interperiodica".*

Acoustic methods are widely used in monitoring hydrobionts for solving different problems, such as tracking the migration of large marine mammals [1] and estimating the amount of anadroms moving to spawning sites. In the context of such problems, some researchers [2, 3] use the sound field perturbation appearing in the stationary acoustic path between an immobile radiator and an immobile receiver because of the sound diffraction by the body of an animal (a scattering body) crossing the path. The use of a monochromatic sound source is a characteristic feature of these studies (see also review [4]). This paper considers the possibility of using wideband radiation in such monitoring problems. The use of wideband radiation may extend the potentialities of the method and, in some cases, enhance its efficiency.

Figure 1 schematically represents the problem under investigation. An immobile radiator is located at point *A*, and an immobile receiver, at point *B*. A scattering body of length *d* moves along the straight line with constant velocity *V* so as to cross the *AB* line at point *C* (*AC* = *CB* = *R*) at the instant of time *t* = 0. We use these geometrical restrictions to simplify further calculations. For the same purpose, we consider a free space without taking into consideration the waveguide effects; in addition, we neglect the spherical divergence of the signal and the attenuation. The radiator located at point *A* generates a sound field $p_0\left(t + \frac{2R}{c}\right)$ in the frequency band $\Delta\omega$, from ω_1 to ω_2 (the central frequency is $\omega_0 = (\omega_1 + \omega_2)/2$). We assume that the Fraunhofer field condition $R \gg d^2/\lambda$ holds for all frequencies of the signal band.

The sound wave is incident on the scattering body at angle φ, and the receiver receives the field scattered at angle $-\varphi$. Because the direct and scattered waves travel along paths of different geometrical lengths, the scattered wave is delayed by $\Delta t = 2(\sqrt{L^2 + R^2} - R)/c$, where $L = Vt$ is the current distance between the body and point *C*. We consider only small angles φ; in this case, $\Delta t = V^2t^2/Rc = \alpha t^2$ and $\alpha = V^2/Rc$. Thus, for small angles ϕ, the delay of the scattered sound wave relative to the

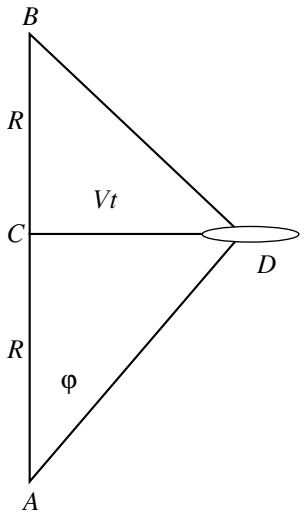


Fig. 1. Schematic diagram of the detection in transmitted sound.

1063-7710/04/5003-0318$26.00

direct wave is a quadratic function of time up to the moment at which the body reaches line *AB*. With such a behavior of the delay, the scattered signal acquires a linear frequency modulation in the case of an insonification with a tone signal, and this fact is used in signal processing [2–4].

To describe the scattered field at the reception point, one must consider the problem of diffraction by the scattering body. The common practice consists of modeling the scattering body with a certain surface (e.g., a spheroid [2]) that allows a separation of variables in the Helmholtz equation. In the case of wideband signals under consideration, such an approach will require solving the problem for all frequencies of the frequency band $[\omega_1, \omega_2]$, which leads to time-consuming computations.

We solve the diffraction problem using the alternative approach developed by M.V. Fedoryuk [5, 6]. For simplicity, we consider opaque bodies. It is known that, in the first approximation, only the shadow contour determines the field scattered in the forward direction (in the shadow region) by an opaque body that is large in comparison with the wavelength, and the effect of other characteristics of the body is negligible. This property holds to a certain extent even for lower frequencies. In the cited papers, M.V. Fedoryuk obtained an approximate solution to the problem of diffraction by bodies of revolution whose longitudinal size was assumed to be large in comparison with the wavelength ($kd \gg 1$). An asymptotic representation of the directivity pattern in the far field zone was obtained by sewing together with the long-wave asymptotic expression for the near field in a certain intermediate zone. In the context of our consideration, the essential point is that the shadow-forming lobe corresponds to the above short-wave asymptotics. The expression for this lobe is simply the integral of the radiation of imaginary secondary sources that are distributed along the axis of revolution according to a certain (depending on the shape of the body) density and that have amplitudes proportional to the incident field. If requirements imposed on the size of the body in comparison with the wavelength hold for all frequencies within $[\omega_1, \omega_2]$, the corresponding procedure of calculating the diffracted field appears to be applicable not only to monochromatic but also to wideband signals.

We use these results to evaluate the scattered field. On the segment $x \in [L - d/2, L + d/2]$, the direct field of the radiator has the form

$$p = p_0\left(t + \frac{2R}{c} - \frac{\sqrt{(L+x)^2 + R^2}}{c}\right).$$

Then, the scattered field at the reception point is

$$p_1 = \frac{\gamma}{d}\int_{-\frac{d}{2}}^{\frac{d}{2}} \mu(x)\, p_0\left(t + \frac{2R}{c} - \frac{2\sqrt{(L+x)^2 + R^2}}{c}\right)dx = \frac{\gamma}{d}\int_{-\frac{d}{2}}^{\frac{d}{2}} \mu(x)\, p_0\left(t - \alpha t^2 - \frac{2Lx}{Rc}\right)dx, \quad (1)$$

where γ is a constant whose value depends on the scattering cross section of the body and on the geometry of the problem, and $\mu(x)$ is the density of secondary sources. Below, we will assume this density to be constant $\mu(x) = 1$, which corresponds to a cylindrical body (it is assumed that the nonsmooth edges only slightly contribute to the scattered field). In this case, integral (1) has a simple meaning: for any monochromatic component $\exp(-i\omega t)$ of field p_0, the integration over x gives the factor $\sin(kd\varphi)/kd\varphi$ (where $\varphi = L/Rc$) that is typical of the diffraction by a strip of width d.

In what follows, we will denote the current time of an observer at the reception point as t' and the time required for the scattering body to travel the distance L to line *AB* as t.

Formula (1) allows us to construct a model of the desired signal from the received direct field. The simplest case of such a construction is realized for small distances L (near the center of the shadow-forming lobe of the directivity pattern). More precisely, we can neglect the quantity $2Lx/Rc$ if the condition $Ld/Rc \ll 2\pi/\omega$ is satisfied for all frequencies of the band, and Eq. (1) at $\mu(x) = 1$ takes the form

$$p_1(t') \approx \gamma p_0(t' - \alpha t^2). \quad (1')$$

This means that the scattered field replicates the direct field, but has a smaller amplitude and is delayed in time by the difference between the times required for the sound wave propagation through distances *AB* and *ADB* (Fig. 1). Neglecting the quantity $2Lx/Rc$ is equivalent to replacing the main lobe of the directivity pattern, $\sin(kd\varphi)/dk\varphi$, with a peak value equal to unity. More accurate expressions for the diffracted field can be obtained in a number of ways. In particular, one can expand the integrand in the Taylor series in $2Lx/Rc$ and retain as many terms as required for a sufficiently accurate description of the main lobe of the function $\sin(kd\varphi)/kd\varphi$; in the case of an even function $\mu(x)$, the expansion will contain only even derivatives. In this case, the integral allows an explicit calculation. In what follows, we will use only the first approximation, because further corrections do not affect the essence of the processes.

The total field at the reception point has the form

$$p(t') = p_0(t') + p_1(t') \approx p_0(t') + \gamma p_0(t' - \alpha t^2), \quad (2)$$

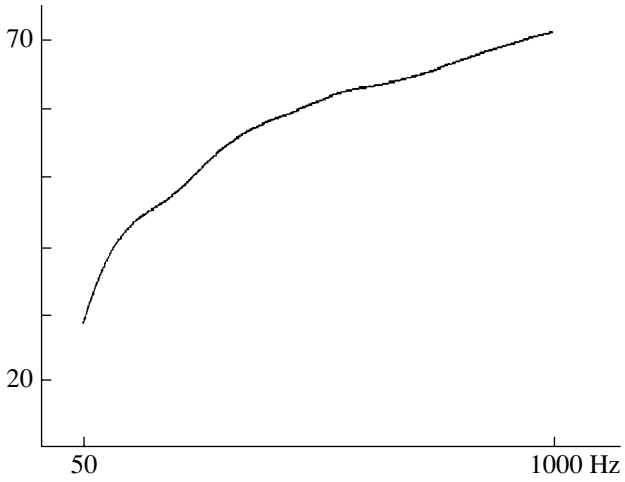


Fig. 2. Output signal-to-noise ratio as a function of signal bandwidth Δf. The accumulation during $T = 2$ s.

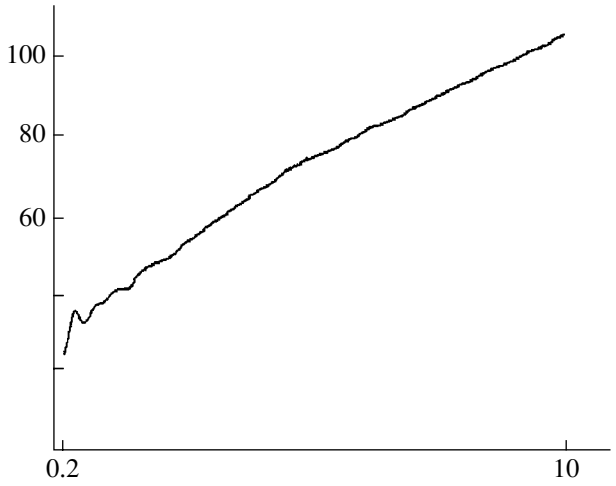


Fig. 3. Output signal-to-noise ratio as a function of the accumulation duration T. The signal bandwidth is Δf = 400 Hz.

and the scattered field is much smaller than the direct field, so that $p(t') \approx p_0(t')$. The right-hand side of Eq. (2) is valid only for times t at which the receiver appears within the main lobe of the directivity pattern (with the above reservations).

The desired signal is extracted as follows. We hypothesize that the model of the desired signal coincides with the received direct field at the instant t' at which the body crosses line AB (and which corresponds to $t = 0$) and that the signals are accumulated within a certain temporal interval (symmetric about the instant $t = 0$), whose duration T depends on the width of the main lobe of the directivity pattern. If the hypothesis is valid, the resulting signal will have a peak characteristic of the cross-correlation processing at the instant corresponding to $t = 0$ and will randomly oscillate for other instants. Hence, the processing consists in the calculation of the integral

$$I(t') = \int_{-\frac{T}{2}}^{\frac{T}{2}} p(t+t')\,p(t+t'-\alpha t^2)\,dt. \qquad (3)$$

If the hypothesis is valid, we can replace the first term of the integrand in accordance with Eq. (2), and the second term, with $p_0(t + t' - \alpha t^2)$, due to the smallness of parameter γ. Then, the integral takes on the form

$$\begin{aligned} I(t') &= I_1(t') + I_2(t') \\ &= \int_{-\frac{T}{2}}^{\frac{T}{2}} p_0(t+t')\,p_0(t+t'-\alpha t^2)\,dt \\ &\quad + \gamma \int_{-\frac{T}{2}}^{\frac{T}{2}} (p_0(t+t'-\alpha t^2))^2\,dt. \end{aligned} \qquad (4)$$

Consider this integral for the case of process $p_0(t)$ specified as white noise filtered in the frequency band $\Delta\omega$. The second integral depends on the presence of the scattering body and forms the signal peak of the output process. The integration accumulates the direct signal, so that the peak value estimator is $|I_2(t')| \approx \gamma\overline{(p_0(t'))^2}\,T$. Let us estimate the width of this peak. If the hypothetic instant t at which the body crosses line AB is delayed by τ, the delay of the model signal will remain equal to αt^2, whereas the delay of the actual signal will be $\alpha(t-\tau)^2$ and the difference will be $\alpha t^2 - \alpha(t-\tau)^2 \approx 2\alpha t\tau$. For the phase difference between the model and actual factors in integral $I_2(t')$ to be acceptable, it is necessary that the condition $2\alpha t\tau < \pi/\omega_0$ be satisfied at the end of the integration interval. From this condition at $t = T/2$, we obtain $\tau < \pi/\alpha T\omega_0 = R\lambda_0/2V^2T$, where λ_0 is the wavelength at the central frequency. Taking into account the fact that the maximum angle $\varphi = VT/2R$ (the zero of the directivity pattern) is equal to $\lambda_0/2d$ at frequency ω_0, we obtain $T = R\lambda_0/Vd$, which yields the estimate $\tau < d/2V$. Thus, the total width of the signal peak, $\Delta t'$, can be estimated as the time d/V required for the body to cross the sound propagation path.

In the first integral, which is the interference caused by the direct signal, the factors are shifted in time and appear noncoherent for times $t > \sqrt{1/\alpha\Delta\omega}$ (the central segment corresponding to small shifts can be "cut out" from the integration). For this reason, low-frequency fluctuations of the integral are increased proportionally to $\overline{(p_0(t'))^2}\sqrt{T/\Delta\omega}$ with increasing duration of the accumulation.

If the hypothesis that the body crosses line *AB* at instant t' fails, the second term will disappear from Eq. (2) for the received signal and, correspondingly, no integral $I_2(t')$ describing the coherent accumulation will appear.

Thus, when the hypothesis is valid, the initial signal-to-noise ratio increases by approximately a factor of $\sqrt{\Delta\omega T}$ and the signal peak value is proportional to the accumulation duration and is linear in parameter γ.

To check the above estimates, we simulated the detection process on the basis of calculating integral (3). In simulations, we used the pseudonoise signal with a carrier frequency $f_0 = \omega_0/2\pi = 1000$ Hz and a bandwidth $\Delta f = \Delta\omega/2\pi$ varying from 20 to 1000 Hz (these particular values are fairly arbitrary; only the product $\Delta\omega T$ and the maximum value of angle φ in Fig. 1 are significant). This signal was used to form, in accordance with Eq. (1'), the model of the diffracted signal with a maximum duration $T = 10$ s and a small signal-to-noise ratio $\gamma = 0.01$–0.1. We superimposed the short model signal on the long initial signal at the center of the realization and then processed the product according to Eq. (3). From the calculated results, we determined the ratio of signal peak value to the root-mean-square value of interference fluctuation.

Figure 2 shows the resulting signal-to-noise ratio versus bandwidth Δf at the accumulation duration $T = 2$ s (on the linear scale). Figure 3 shows this quantity as a function of accumulation duration T for bandwidth $\Delta f = 400$ Hz. The results obtained agree well with the above estimates.

Experimentally, we checked the efficiency of the proposed method in a shallow-water gulf with depths of 30–40 m under the conditions of a homogeneous waveguide in the presence of a small surface roughness. The quasi-stationary acoustic path between the cutter moored to the anchorage buoy and the vessel on the berth that received signals with the use of its regular acoustic system had a length of 600 m. An omnidirectional radiator of the pseudonoise insonification signal in the frequency band from 3 to 7 kHz was lowered from the deck of the cutter down to a depth of 10 m. The vessel received signals at a depth of about 8 m using its regular acoustic system. The objects to be detected were scuba divers who traversed the path at its center in the perpendicular direction at a depth of 10 m. The scuba divers were dressed in hydrosuites and were equipped with a towing unit, which increased the effective size of the scattering volume. The longitudinal size of this volume was roughly estimated to be about 2 m. The speed of motion of a scuba diver was approximately 1 m/s.

Experimental estimates show that the insonification signal exceeded the sea noise by no less than 15 dB.

The preliminary parameters of processing $\alpha = 2.2 \times 10^{-6}$ s^{-1} and $T = 45$ s were obtained from the experimental geometry and central frequency of the insonification signal. They were improved in the course of the signal processing performed in accordance with Eq. (3). The maximum signal peak was obtained for $\alpha = 1.53 \times 10^{-6}$ s^{-1} and $T = 60$ s. This value of parameter α corresponds to the scatterer speed $V = 0.83$ m/s. With this speed, the theoretical estimate of the accumulation duration measures 56.25 s, which is close to the value obtained in the course of processing.

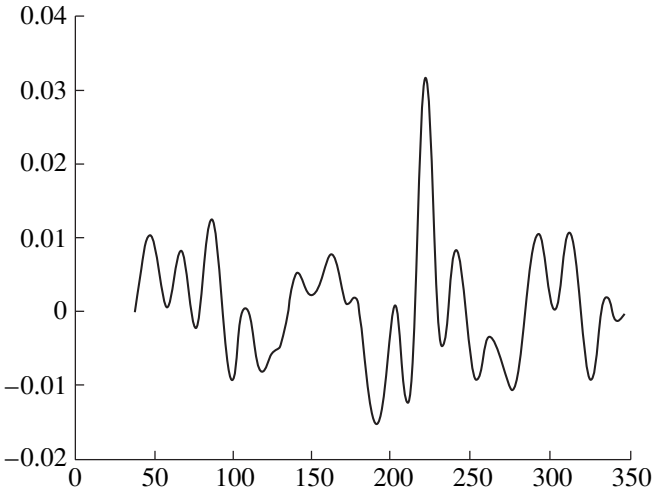


Fig. 4. Output of processing according to Eq. (3).

Figure 4 shows one of the processing outputs obtained in the case of traversing the sound path by a scuba diver. We rejected the low-frequency components of this signal by using the fourth-order Butterworth filter with a cutoff frequency of 8.3×10^{-3} Hz (the corresponding period is 120 s). Time (in seconds) measured from the start of record is plotted on the horizontal axis, and the processing output is plotted on the vertical axis. At the 221st second corresponding to traversing the acoustic path by a scuba diver, a sharp peak noticeably exceeding the background fluctuation of the output is observed. The output signal-to-noise ratio calculated as the ratio of the peak value to the root-mean-square deviation of fluctuation measures 13.8 dB, which is quite sufficient for detecting the scatterer.

It should be noted that experimental conditions essentially differ from the idealized representation considered in this paper. First of all, the experiment was carried out under the conditions of multipath propagation, whose effect on the processing procedure suggested here requires special consideration, which is beyond the scope of this paper. One can expect that, in a multipath channel, the efficiency of processing will be lowered as in the case of traditional correlation processing, so that the result obtained is not necessarily the maximum possible one.

Another difficulty consists in the appearance of an additional type of moving scatterers—floating bubbles of air expired periodically by the scuba diver. Although the speed of bubbles (0.1–0.3 m/s) is less than the speed

of the scuba diver, they scatter sound very strongly and may substantially contribute to the background fluctuation of the processing output even being located at large distances from the acoustic path.

Thus, the experiment showed that the method considered here appears to be efficient even under relatively complex conditions and, hence, offers considerable promise in solving the problems of monitoring hydrobionts.

REFERENCES

1. J. Gedamke, D. Costa, and A. Dunstan, J. Acoust. Soc. Am. **109**, 3038 (2001).
2. V. M. Bel'kovich, V. A. Grigor'ev, B. G. Katsnel'son, and V. G. Petnikov, Akust. Zh. **48**, 162 (2002) [Acoust. Phys. **48**, 133 (2002)].
3. V. A. Grigor'ev, B. G. Katsnel'son, V. M. Kuz'kin, and V. G. Petnikov, Akust. Zh. **47**, 44 (2001) [Acoust. Phys. **47**, 35 (2001)].
4. S. M. Gorskiĭ, V. A. Zverev, and A. I. Khil'ko, in *Formation of Acoustic Fields in Oceanic Waveguides*, Ed. by V. A. Zverev (Inst. Prikl. Fiz., Akad. Nauk SSSR, Nizhni Novgorod, 1991), pp. 82–114.
5. M. V. Fedoryuk, Izv. Akad. Nauk SSSR, Ser. Mat. **49** (1), 160 (1985).
6. M. V. Fedoryuk, Akust. Zh. **27**, 605 (1981) [Sov. Phys. Acoust. **27**, 336 (1981)].

Translated by A. Vinogradov

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 323–330. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 389–396.

BIOLOGICAL ACOUSTICS

Resonances of a Branched Vocal Tract with Compliant Walls

I. S. Makarov and V. N. Sorokin

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

e-mail: speechprod_mak@mail.ru

Received January 14, 2003

Abstract—The calculation of the resonance frequencies from experimental cross-sectional areas of a vocal tract under the assumption that its walls are perfectly rigid provides values that noticeably differ from the measured resonance frequencies. The compliance of the walls affects the first resonance and almost does not affect the higher-order resonances. The presence of branching in the tract at the level of the larynx affects the second and third resonances stronger than the first resonance. The parameters of the wall impedance (the loss, mass, and elasticity) and the length and cross-sectional area of the branchings are determined by minimizing the rms discrepancy between the measured and calculated resonance frequencies. The error in the frequency calculation with allowance for the wall compliance and branching in the tract proves to be within the accuracy of the formant estimation. © *2004 MAIK "Nauka/Interperiodica".*

In articulatory speech synthesis, a problem often arises of evaluating the resonance frequencies of a vocal tract from its cross-sectional area given as a function of the longitudinal coordinate [1]. When solving an inverse problem for the cross-sectional area or shape of the vocal tract for vowels, the acoustic data are given in the form of measured formant frequencies [2]. The resonance frequencies of a vocal tract with compliant walls are usually calculated as a solution to the spectral problem with boundary conditions for the wave equation

$$\frac{\partial}{\partial x}\left[S(x)\frac{\partial p}{\partial x}\right] = \frac{S(x)}{c_0^2(x)}\frac{\partial^2 p}{\partial t^2}, \quad (1)$$

where $S(x)$ is the cross-sectional area of the vocal tract, x is the coordinate along the tract's midline, $p(x, t)$ is the acoustic pressure, and $c_0(x)$ is the velocity at which acoustic waves propagate in the tract. This equation is valid for frequencies below 4.5 kHz and a time-independent area $S(x)$ [3].

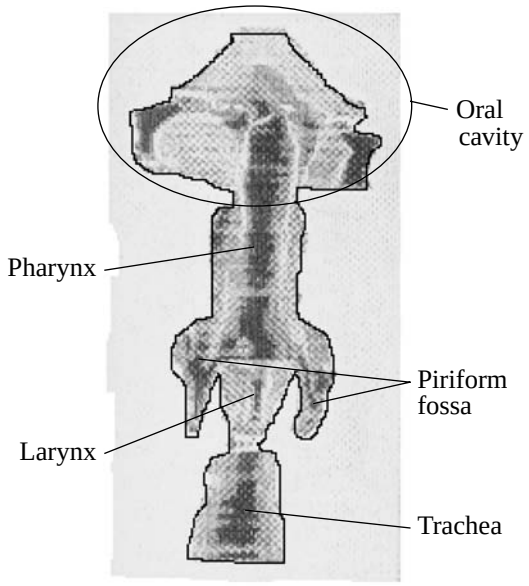
Before methods for the direct measurement of the cross-sectional area of a vocal tract were developed, various methods for calculating the resonance frequencies from the cross-sectional area were addressed [4, 5]. A disagreement between measured resonance frequencies of the vocal tract and those calculated from the tract's area measured by magnetic resonance imaging (MRI) was demonstrated in [6–8].

It is known that the impedance of the tract's walls of an inertial or compliant type may increase or decrease the first resonance frequency [5]. The lower part of the vocal tract contains piriform branchings (called piriform fossa or sinus piriformis), which have the form of two conical cavities connected to the vocal tract on both sides of the pharynx approximately at the level of entrance to the esophagus (Fig. 1). Piriform cavities may considerably affect the resonance frequencies of the vocal tract [7, 9]. However, results reported in different works are contradictory. In particular, piriform fossa were reported to exert the maximum effect on the first resonance (the first resonance frequency may be reduced by 6 to 10%), while the effect on the higher resonances is comparatively small (the higher resonance frequencies may be reduced by 1.2 to 3%) [9]. On the contrary, paper [7] indicates that piriform fossa exert the maximum effect on the second and third resonance frequencies, while the effect on the first resonance frequency is weaker.

Thus, it is necessary to study the factors that affect the resonance frequencies of the vocal tract; i.e., it is necessary to develop a more realistic acoustic model.

At present, the main experimental technique for measuring the cross-sectional area of the vocal tract is magnetic resonance imaging (MRI) [7]. MRI is a technique that creates tomographic images of a human body in terms of magnetic resonance. When the scanning is performed, a number of plates are put against the speaker's head so that they are positioned across the vocal tract. The speaker is placed inside a magnet, which creates a magnetic field. The magnetic field is necessary to rotate the hydrogen spins by 90° or 180°. Placed inside the magnet are RF coils, which excite the hydrogen spins in each layer by RF pulses and detect the reflected NMR signal. This signal is subjected to a two-dimensional Fourier transform, which creates the magnetic resonance image of each layer [8]. During the scanning, which usually lasts for 5 to 15 s, the speaker must keep the articulation unchanged. Magnetic resonance tomography, unlike X-ray measurements used

1063-7710/04/5003-0323$26.00



Vocal tract: rear view.

earlier, does not impose any limitation on the duration of the experiment in terms of the radiation dose.

Because magnetic resonance studies are difficult to perform, there are very little experimental data available on the cross-sectional areas of the vocal tract. As a rule, these data refer to stationary vowels.

One of the main methods for evaluating the resonance frequencies is linear prediction. The systematic error in estimated formant frequencies associated with pitch was shown to be about 10% of the pitch [10]. Systematic errors associated with the order of the model, with the method used to find the poles, and when the number of a specific pole are within 10 to 80 Hz. Thus, the linear prediction evaluates the formant frequencies to within 8 to 10%.

The MRI data on the cross-sectional areas of vowels and the first three resonance frequencies for each vowel were borrowed from [6, 7]. Two sets of area functions were used. Each set consisted of 10 area functions recorded during the articulation of 10 English vowels: /i/ (heat), /ɪ/ (hit), /ɛ/ (head), /æ/ (hat), /Λ/ (ton), /α/ (hot), /⊃/ (paw), /o/ (hoe), /∪/(hood), and /u/(who). The first data set refers to the vocal tract of a male speaker, the second one, to a female speaker. The experiments used 24 plates 5 mm thick each, which were spaced 5 mm apart.

In these publications, the resonance frequencies were calculated over a 25-ms-long Hamming window by the linear prediction autocorrelation method using a 16th-order model. The poles of the transfer function were calculated by finding peaks of the Fourier spectrum with a further parabolic interpolation. Along with the measured resonance frequencies, the authors of [6, 7] used the measured cross-sectional areas to evaluate the resonance frequencies by the modified traveling wave scheme [11, 12]. This scheme allowed for the wall compliance and the impedance due to the radiation through the lips. The branching was ignored in [6], while in [7] it was taken into account as an additional area in the pharynx.

The speech production theory employs several models for evaluating the resonance frequencies from the cross-sectional areas [1]. The acoustic model used in our study relies on the transmission-line method. It allows for the wall compliance and for the branching in the tract. The model represents the vocal tract as a combination of N tubes of constant cross-sectional areas. The pressure P_l and volume velocity U_l at the lips are related to the pressure P_g and volume velocity U_g at the glottis by the formula

$$\begin{pmatrix} P_g \\ U_g \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} P_l \\ U_l \end{pmatrix}, \tag{2}$$

where

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \prod_{i=1}^{N} \begin{pmatrix} \cosh(\gamma l_i) & \frac{\rho \bar{c}_0}{S(i)} \sinh(\gamma l_i) \\ \frac{S(i)}{\rho \bar{c}_0} \sinh(\gamma l_i) & \cosh(\gamma l_i) \end{pmatrix}. \tag{3}$$

Here, N is the total number of the tubes, $S(i)$ is the cross-sectional area of the ith tube, l_i is the length of the ith tube, $\gamma = \beta + jk$, β is the attenuation coefficient, $k = \omega/\bar{c}_0$ is the wave number, ω is the circular frequency (in radians per second), ρ is the air density (0.00114 g/cm^3), $\bar{c}_0$ is the sound velocity in the tube with hard walls (35000 cm/s), and j is the imaginary unit. The vocal tract's transfer function is defined as the ratio of the volume velocity at the lips to that at the glottis, i.e., $H(\gamma) = U_l/U_g$. The input impedance, as seen looking into the glottis, is a pressure-to-volume velocity ratio at the glottis, i.e., $Z_g = P_g/U_g$. From formulas (2) and (3), we obtain

$$H(\gamma) = \frac{1}{CZ_l + D}, \tag{4}$$

$$Z_g = \frac{AZ_l + B}{CZ_l + D}. \tag{5}$$

Here, Z_l is the radiation impedance at the lips defined as $Z_l = \rho \frac{(2\pi F)^2}{2c_0} + j\frac{16F\rho}{3R_0}$, where R_0 is the equivalent radius of the lips' aperture and F is the frequency [13].

As follows from Eqs. (4) and (5), the poles of the transfer function coincide with those of the input impedance. Therefore, to find the resonance frequencies, it is sufficient to calculate the transfer function or the input impedance by the transmission-line method. In this study, we calculate the input impedance.

Table 1. Resonance frequencies of the vocal tract with rigid walls without piriform regions: a male speaker (Titze, 1996)

Vowel	F1 (meas.)	F1 (calc.)	Error (%)	F2 (meas.)	F2 (calc.)	Error (%)	F3 (meas.)	F3 (calc.)	Error (%)
/i/ heat	333	226.5	–31.9	2332	2575	10.4	2986	3652	22.3
/ɪ/ hit	518	487.6	–5.9	2004	2141	6.9	2605	2729	4.8
/ɛ/ head	624	646.9	3.7	1853	2064	11.4	2475	2665	7.8
/æ/ hat	692	765.9	10.7	1873	1952	4.2	2463	2665	8.2
/ʌ/ ton	707	753.8	6.7	1161	1363	17.4	2591	2679	3.4
/ɑ/ hot	754	808	7.2	1195	1209	1.2	2685	2895	7.8
/ɔ/ paw	654	644.9	–1.4	944	1040	10.2	2739	2213	–19.2
/o/ hoe	541	382.8	–29.2	1045	926	–11.4	2568	2721	5.9
/ʊ/ hood	540	395	–26.8	922	893	–3.1	2584	2547	–1.4
/u/ who	389	261.5	–32.8	987	1206	22.2	2299	2480	7.9

Table 2. Resonance frequencies of the vocal tract with rigid walls without piriform regions: a female speaker (Titze, 1998)

Vowel	F1 (meas.)	F1 (calc.)	Error (%)	F2 (meas.)	F2 (calc.)	Error (%)	F3 (meas.)	F3 (calc.)	Error (%)
/i/ heat	489	316.6	–35.3	2387	2866	20.0	3526	3262	–7.5
/ɪ/ hit	649	714	10.0	2079	2328	11.9	2799	3511	25.4
/ɛ/ head	799	856.7	7.2	2112	2312	9.5	2874	3394	18.0
/æ/ hat	837	1016	21.4	2028	2536	25.0	2814	4042	43.6
/ʌ/ ton	759	903.6	19.0	1360	1604	17.9	2756	3577	29.8
/ɑ/ hot	961	828.4	–13.8	1488	1644	10.5	2779	3058	10.0
/ɔ/ paw	817	819.9	1.0	1264	1369	8.3	2574	2963	15.1
/o/ hoe	706	777.6	10.2	1157	1588	37.3	3030	3281	8.3
/ʊ/ hood	675	920.9	36.4	1122	1616	44.0	3157	3098	–1.9
/u/ who	533	458.5	–13.9	1152	1233	7.0	3122	3162	1.3

The results of calculating the resonance frequencies are shown in Tables 1 and 2. From these tables, one can see that the resonance frequencies calculated by the transmission-line method from the experimental cross-sectional areas under the assumption that the walls are rigid were in many cases noticeably different from the measured resonance frequencies. The error was often greater than the error within which the formants are estimated.

For the first resonance, this manifests itself in the vowels /i/ (heat), /æ/ (hat), /o/ (hoe), /u/ (hood), and /u/ (who) of both speakers and in the vowels /ɪ/ (hit), /ʌ/ (ton), and /ɑ/ (hot) of the female speaker. For the second resonance, the greatest errors occur in the vowels /i/ (heat), /o/ (hoe), and /ʌ/ (ton) of both speakers; /ɛ/ (head), /ɔ/ (paw), and /u/ (who) of the male speaker; and /ɪ/ (hit), /æ/ (hat), and /ʊ/ (hood) of the female speaker. For the third resonance, the greatest errors are observed in the vowel /ɔ/ (paw) of both speakers; the vowel /i/ (heat) of the male speaker; and the vowels /ɪ/ (hit), /ɛ/ (head), /æ/ (hat), and /ʌ/ (ton) of the female speaker.

The frequencies obtained by this algorithm were compared with those evaluated by other methods, namely, by the finite-difference scheme applied to the speech signal, as described in [1], and by the technique that locates the poles of the transfer function calculated with the help of the traveling wave method [14]. Resonance frequencies calculated by these methods and those calculated in [6, 7] were found to be close to each other. Therefore, the disagreement observed between the theoretical and experimental resonance frequencies should be attributed to the model of the vocal tract rather than to the algorithm.

As seen from Tables 1 and 2, the second and third calculated resonances are usually higher than the second and third measured resonances. As for the first calculated resonance, for the male speaker it may be either higher than the first measured resonance without going beyond the accuracy of formant estimation or much lower than the first measured resonance. For the female speaker, the first calculated resonance may be much lower or much higher than the first measured resonance. Since the greatest effect on the first resonance is exerted by the compliance of the walls, we studied the role played by the parameters of the wall impedance. When the walls of the vocal tract are compliant, the column of air in the tract becomes smaller under the action

of pressure not only due to the compressibility of the medium but also because of the wall displacement. It has been shown [15] that the velocity of sound in a tube with compliant walls is given by the expression

$$c_0^2 = \frac{\bar{c}_0^2}{1 - \dfrac{2\rho\bar{c}_0^2}{j\omega Z_\omega a}}, \quad (6)$$

where Z_ω is the impedance, a is the characteristic transverse size of the tube, and $\bar{c}_0$ is the velocity of sound in the tube with perfectly rigid walls. The thickness of the tube's walls imposes limitations on the applicability region of formula (6), because, if the thickness is comparable with the wavelength of the radial wave, the properties of the walls are not completely described by their impedance: they also depend on the distributions of the acoustic pressure and particle velocity along the tube [16]. The walls of the vocal tract can be considered as thin in terms of the wavelength of the radial oscillations, and formula (6) can be applied to estimate the acoustic processes in the vocal tract. The wall impedance Z_ω can be written in the general form as

$$Z_\omega = r + j\left(\frac{c}{\omega} - \omega m\right), \quad (7)$$

where r is the loss in the walls, m is the mass of the walls, and c is their compliance.

As can be seen from (6), the wall compliance affects the velocity of sound at relatively low frequencies [5, 17]. The direct use of relationship (6) in the transmission-line method to calculate the resonance frequencies encounters a number of problems. When the geometric parameter a is an arbitrary function of the x coordinate, we obtain a system of nonlinear equations, which is difficult to solve. Therefore, we used the modified transmission-line procedure proposed in [18], which takes into account the parameters of the wall impedance. This procedure uses Eq. (2) and replaces Eq. (3) with the relationship

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \prod_{i=1}^{N} \begin{pmatrix} \cosh\left(\dfrac{\sigma l_i}{\bar{c}_0}\right) & \dfrac{\rho\bar{c}_0}{S(i)}\varphi\sinh\left(\dfrac{\sigma l_i}{\bar{c}_0}\right) \\ \dfrac{S(i)}{\rho\bar{c}_0\varphi}\sinh\left(\dfrac{\sigma l_i}{\bar{c}_0}\right) & \cosh\left(\dfrac{\sigma l_i}{\bar{c}_0}\right) \end{pmatrix}. \quad (8)$$

Here, l_i is the length of the ith tube, $\varphi = -\sqrt{\omega}/\sqrt{\theta + j\omega}$, ω is the circular frequency, $\sigma = \varphi(\theta + j\omega)$, $\theta = \dfrac{j2.6\bar{c}_0\pi^2\omega}{j\omega(1+\chi)+\delta}$, $\chi = \dfrac{r}{m}$, and $\delta = \dfrac{1}{mc}$, where r is the loss in the walls, m is the mass of the walls, and c is the wall's elasticity.

The tissues of the vocal tract's walls have different characteristics. The tongue, lips, cheeks, and soft palate are more or less compliant, while the hard palate consists of a bone base covered by a comparatively thin layer of soft tissue. It is therefore natural to assume that the parameters of the wall impedance (the loss, mass, and elasticity) are functions of the coordinate along the tract.

The parameters of the wall impedance were found for each speaker by minimizing the rms discrepancy between the calculated and measured values of the first resonance frequency. The loss, mass, and elasticity, which vary along the tract, were represented by the first eight terms of the Fourier series

$$r(x) = R_0 + \sum_{\xi=1}^{4} r_{1\xi}\sin\left(\frac{\xi\pi x}{l}\right) + \sum_{k=1}^{4} r_{2\xi}\cos\left(\frac{\xi\pi x}{l}\right),$$

$$m(x) = M_0 + \sum_{\xi=1}^{4} m_{1\xi}\sin\left(\frac{\xi\pi x}{l}\right) + \sum_{k=1}^{4} m_{2\xi}\cos\left(\frac{\xi\pi x}{l}\right),$$

$$c(x) = C_0 + \sum_{\xi=1}^{4} c_{1\xi}\sin\left(\frac{\xi\pi x}{l}\right) + \sum_{\xi=1}^{4} c_{2\xi}\cos\left(\frac{\xi\pi x}{l}\right).$$

Here, l is the tract's length. The parameters r, m, and c were optimized for each speaker over all vowels by minimizing the functional

$$\Phi(z) = \sum_{n=1}^{T} [F_m^{(1n)} - F_c^{(1n)}(z)]^2 \quad (9)$$

under the constraints $R = \{r, r1 \le \min(r); \max(r) \le r2\}$, $M = \{m, m1 \le \min(m); \max(m) \le m2\}$, and $C = \{c, c1 \le \min(c); \max(c) \le c2\}$. Here, z is the vector with the components (r, m, c), $T = 10$ is the number of vowels involved in the experiment, $F_m^{(1n)}$ is the measured first resonance frequency for the nth sound, and $F_c^{(1n)}$ is the calculated first resonance frequency for the nth sound. According to various sources, the loss lies between 540 and 11 000 g/s cm^2; the mass, between 1.08 and 2.07 g/cm^2; and the elasticity, between 10^3 and 10^6 Pa/cm. Hence, the constraints for R, M, and C were specified as $R = \{r, 540 \le r \le 11000$ g/s cm$^2\}$, $M = \{m, 1.08 \le m \le 2.7$ g/cm$^2\}$, and $C = \{c, 10^3 \le c \le 10^6$ Pa/cm$\}$ [5]. Along with the variable wall parameters, we studied the effect of the effective (constant) parameters by minimizing functional (9). The constraints for R, M, and C were specified as in the scheme with the wall impedance parameters varying along the tract.

It was found that taking into account the variation of the wall impedance parameters along the tract actually does not improve the results, as compared to those obtained with constant parameters. In some cases (/ɪ/ (hit), /Λ/ (ton), /ɑ/ (hot), /ʊ/ (hood)) for a female speaker), parameters varying along the tract even lead to inferior results. Therefore, when calculating the first resonance frequency with allowance for compliance, it is sufficient to assume that the parameters are constant along the tract.

Table 3. First resonance without and with taking into account the wall compliance: a male speaker (Titze, 1996). The loss is 542.4 g/s cm^2, the mass is 2.4 g/cm^2, and the elasticity is 1.5×10^3 Pa/cm

Vowel	F1 (meas.)	F1 (calc. without compl.)	Error (%)	F1 (calc. with a constant compl.)	Error (%)	F1 (calc. with a varying compl.)	Error (%)
/i/ h<u>ea</u>t	333	226.5	–31.9	375.9	12.9	375.3	12.7
/l/ h<u>i</u>t	518	487.6	–5.9	551.6	6.5	549.1	6
/ɛ/ h<u>ea</u>d	624	646.9	3.7	635	1.8	625.9	0.3
/æ/ h<u>a</u>t	692	765.9	10.7	691	–1.1	680.2	–1.7
/Λ/ t<u>o</u>n	707	753.8	6.7	685.7	–3.2	684.4	–3.2
/α/ h<u>o</u>t	754	808	7.2	710	–5.8	708.8	–6
/⊃/ p<u>aw</u>	654	644.9	–1.4	634	–3.1	630.7	–3.56
/o/ h<u>oe</u>	541	382.8	–29.2	488.8	–9.6	492.9	–8.9
/∪/ h<u>oo</u>d	540	395	–26.8	497	–8.2	498.9	–7.6
/u/ wh<u>o</u>	389	261.5	–32.8	403.9	3.8	399.9	2.8

Since the wall compliance exerts the greatest effect on the first resonance, we studied how the inclusion of the branching in the consideration refines the second and third resonance frequencies.

The resonance frequencies of the branched tract were evaluated from the poles of the input impedance as follows. At first, the input impedance Z_m of the region from the branch point to the lips was calculated by Eqs. (5) and (8). Then, these formulas were used to calculate the input impedance Z_p of the piriform cavity, as looking from the side of the vocal tract (under the assumption that, on the other side, the piriform cavity abuts upon a perfectly rigid wall). After that, the input impedance Z_in of the region from the branch point to the lips was modified as

$$Z_\mathrm{in} = \frac{Z_\mathrm{m} Z_\mathrm{p}}{Z_\mathrm{m} + Z_\mathrm{p}}. \tag{10}$$

Finally, the tract's input impedance was evaluated as

$$Z_\mathrm{g} = \frac{AZ_\mathrm{in} + B}{CZ_\mathrm{in} + D}. \tag{11}$$

The two piriform branchings can be modeled by two cylindrical tubes of a constant cross section, which shunt the vocal tract at a distance of about 2 cm from the glottis. Computationally, it is preferable to model both branchings by a single equivalent tube. To this end, two models were studied: one model (the three-tube model) allowed for two branchings (in addition to the vocal tract) and the other model (the two-tube model) contained one branching. The three-tube model depends on four parameters: two cross-sectional areas and two tube lengths. The two-tube model is specified by two parameters: the cross-sectional area and the tube length. The parameters of each model were evaluated for each speaker by minimizing the rms discrepancy between the calculated and measured resonance frequencies. In each model, the tract walls were assumed to be perfectly stiff. The discrepancy between frequencies calculated from the three-tube model and the two-tube model were found to be no greater than 0.8%. Therefore, in the following calculations, we used only the two-tube model.

Then, we made an attempt to evaluate the parameters of the wall impedance and the parameters of the branching tube (the length and cross-sectional area) together. The parameters of the wall impedance and the shunting tube were calculated for each speaker by minimizing the functional

$$\Phi(z_1) = \sum_{n=1}^{N} (F_m^{(1n)} - F_c^{(1n)}(z_1))^2 + (F_m^{(2n)} - F_c^{(2n)}(z_1))^2 + (F_m^{(3n)} - F_c^{(3n)}(z_1))^2 \tag{12}$$

under the constraints $R = \{r, r1 \le r \le r2\}$, $M = \{m, m1 \le m \le m2\}$, $C = \{c, c1 \le c \le c2\}$, $S_p = \{s_p, s_p1 \le s_p \le s_p2\}$, and $L_p = \{l_p, l_p1 \le l_p \le l_p2\}$. Here, z_1 is a vector with the components (z, m, c, s_p, l_p), s_p is the cross-sectional area of the branching tube, l_p is the length of this tube, $F_\mathrm{m}^{(\theta n)}$ is the measured resonance frequency, $F_c^{(\theta n)}$ is the calculated resonance frequency, n is the index of the vowel, and c is the index of the resonance frequency. The constraints R, M, and C were defined above. According to [7, 9], the length of the piriform cavity may vary from 1 to 3 cm and the area may be within 0.05 and 1.5 cm^2. Therefore, the S_p and L_p constraints were defined as $S_p = \{s_p, 0.05\ \mathrm{cm}^2 \le s_p \le 1.5\ \mathrm{cm}^2\}$ and $L_p = \{l_p, 1\ \mathrm{cm} \le l_p \le 3\ \mathrm{cm}\}$.

The results of the calculations are summarized in Tables 1 and 2. It can be seen that the simultaneous inclusion of the wall compliance and the branching in the consideration makes the discrepancy between the calculated and measured resonance frequencies fall within the accuracy of estimating the formants in most cases. Sometimes, the inclusion of both compliance of the walls and branchings of the vocal tract in the con-

Table 4. First resonance without and with taking into account the wall compliance: a female speaker (Titze, 1998). The loss is 541.9 g/s · cm^2, the mass is 2.59 g/cm^2, and the elasticity is 1.4×10^5 Pa/cm

Vowel	F1 (meas.)	F1 (calc. without compl.)	Error (%)	F1 (calc. with a constant compl.)	Error (%)	F1 (calc. with a varying compl.)	Error (%)
/i/ h<u>ea</u>t	489	316.6	–35.3	461.8	–5.6	481.2	–1.6
/ɪ/ h<u>i</u>t	649	714	10.0	693.7	6.8	699.6	7.8
/ɛ/ h<u>ea</u>d	799	856.7	7.2	759.6	–4.9	783.8	–1.9
/æ/ h<u>a</u>t	837	1016	21.4	827.5	–1.2	826.9	–1.2
/ʌ/ t<u>o</u>n	759	903.6	19.0	780	2.7	794.7	4.7
/ɑ/ h<u>o</u>t	961	828.4	–13.8	747	–22.3	739.9	–23.0
/ɔ/ p<u>aw</u>	817	819.9	1.0	743.2	–9.0	776.2	–5.0
/o/ h<u>oe</u>	706	777.6	10.2	723.8	2.5	716.6	1.5
/ʊ/ h<u>oo</u>d	675	920.9	36.4	787.6	16.6	789.8	17.0
/u/ wh<u>o</u>	533	458.5	–13.9	555.7	4.2	534.1	0.2

sideration leads to some increase in errors. An example is the vowel /ɔ/ in the word *paw* of the male speaker, where the error in the third resonance increases from –19.2% by 0.3%; however, such an increase in error seems to be insignificant. For a female speaker, the corresponding examples are the vowel /ɑ/ in the word *hot*, where the error in the first resonance increases from –13.8% to –21.3%, and the vowel /ʊ/ in the word *hood*, where the error in the third resonance increases from –1.9% to –16.9%. In some cases, the inclusion of branchings leads to a decrease in the errors in the second and third resonances, but the resulting error values still go beyond the accuracy of formant estimation. For a female speaker, an example is the vowel in the word *hat*, where the error in the third resonance decreases from 43.6% to 27.1%. In one special case (the vowel /u/ in the word *who* of the male speaker), taking branchings into consideration did not affect the second resonance (the error decreased from 22.2% by 0.1%).

In addition to experimental data borrowed from [6, 7], we also used those from [8]. The latter paper reports on the MRI shapes of the vocal tract obtained during the articulation of four vowels (/ɑ/ (h<u>o</u>t), /æ/ (h<u>a</u>t), /i/ (h<u>ea</u>t), and /u/ (wh<u>o</u>)) by two speakers. The experimental technique used in this study was criticized in [6]. Therefore, these data were used as a sort of reference material.

For each vowel, the resonance frequencies were calculated with allowance for the wall compliance and branchings. A simultaneous inclusion of the wall compliance and tract branchings in the consideration reduces the mean errors in all the calculated resonance frequencies by 2 to 5%. At the same time, the error in the second resonance for the male speaker (20.2%) and errors in the first and second resonances for the female speaker (18.1 and 35.7%, respectively) are still greater than the error in estimating the formants. These results can be explained as follows. At present, it is impossible to record the acoustic signal and scan the tract for its shape at the same time. Also, when the tract is scanned by MRI, the speaker has to keep articulation of a vowel unchanged for a rather long time (about 5 to 15 s). A hypothesis advanced in [7, 8] states that these are the very circumstances that are responsible for the considerable disagreement between the measured and calculated resonance frequencies.

The MRI results also depend on the arrangement of the plates used for scanning the vocal tract. Different arrangements of the plates relative to the tract's midline produce different cross-sectional areas, which affects the calculated values of the resonance frequencies. The considerable discrepancies between the calculated and measured resonance frequencies (which were up to 70% for the vowel /u/ (wh<u>o</u>)) were also attributed to a wrong plate positioning [8].

The errors can also be caused by the low resolution of the MRI, which is insufficient for scanning tract regions with a small (less than 0.2 cm^2) cross-sectional areas. It was shown, that, at a plate resolution of 0.0938 cm/pixel, the error between the measured and calculated areas of narrow (with the area of less than 0.2 cm^2) pipes is about 10% [6]. The error increases as the pipe's cross sectional area decreases. It is therefore difficult to estimate the geometric dimensions of narrow branchings in the vocal tract (e.g., piriform fossa) from MRI data.

With the help of MRI, it is also difficult to reconstruct the exact shape of teeth, because their hydrogen content is low [8]. Therefore, the cross-sectional areas of the front part of the vocal tract has to be measured using an artificial palate and dentures fabricated for a particular speaker. This procedure can also be a source of errors in estimating the areas in the hard palate region.

Another possible source of errors is that our study assumed that the parameters of the piriform fossae do not change. However, electron-beam computer tomography shows that the piriform fossae can significantly

Table 5. Resonance frequencies with allowance for the wall compliance and piriform regions: a male speaker (Titze, 1996). The loss is 542.7 g/s · cm^2; the mass is 2.43 g/cm^2; the elasticity is 1.5×10^3 Pa/cm; and the length and area of the piriform branching are 2.1 cm and 0.073 cm^2, respectively

Vowel	F1 (meas.)	F1 (calc.)	Error (%)	F2 (meas.)	F2 (calc.)	Error (%)	F3 (meas.)	F3 (calc.)	Error (%)
/i/ heat	333	377	13.1	2332	2506	7.4	2986	3334	11.6
/ɪ/ hit	518	553	6.7	2004	2095	4.5	2605	2657	2.0
/ɛ/ head	624	636	1.9	1853	1996	7.7	2475	2623	5.9
/æ/ hat	692	693	0.1	1873	1885	1.0	2463	2544	3.3
/ʌ/ ton	707	685	–3.1	1161	1333	14.8	2591	2590	–0.1
/ɑ/ hot	754	704	–6.6	1195	1198	–0.1	2685	2872	6.9
/ɔ/ paw	654	634	–3.1	944	1021	8.1	2739	2205	–19.5
/o/ hoe	541	489	–9.4	1045	917	–12.2	2568	2695	4.9
/ʊ/ hood	540	497	–7.9	922	889	–3.5	2584	2459	–4.8
/u/ who	389	405	4.1	987	1206	22.1	2299	2448	6.5

Table 6. Resonance frequencies with allowance for the wall compliance and piriform regions: a female speaker (Titze, 1998). The loss is 548.1 g/s · cm^2; the mass is 2.7 g/cm^2; the elasticity is 1.5×10^3 Pa/cm; and the length and area of the piriform branching are 2.0 cm and 0.72 cm^2, respectively

Vowel	F1 (meas.)	F1 (calc.)	Error (%)	F2 (meas.)	F2 (calc.)	Error (%)	F3 (meas.)	F3 (calc.)	Error (%)
/i/ heat	489	481	–1.6	2387	2596	8.7	3526	3097	–12.2
/ɪ/ hit	649	703	8.3	2079	1930	–7.1	2799	3013	7.6
/ɛ/ head	799	753	–5.7	2112	1817	–13.9	2874	3017	4.9
/æ/ hat	837	846	1.0	2028	2107	3.9	2814	3577	27.1
/ʌ/ ton	759	742	–2.2	1360	1391	2.3	2756	3133	13.7
/ɑ/ hot	961	756	–21.3	1488	1363	–8.4	2779	2916	4.9
/ɔ/ paw	817	754	–7.7	1264	1139	–9.9	2574	2853	10.8
/o/ hoe	706	737	4.5	1157	1256	8.6	3030	2959	–2.4
/ʊ/ hood	675	734	8.8	1122	1282	14.3	3157	2625	–16.9
/u/ who	533	576	8.1	1152	1195	3.8	3122	2752	–11.9

change their length and volume in the process of articulation [7]. It may also be assumed that a better agreement between the measured and calculated resonance frequencies can be achieved by varying the length and area of the piriform fossae. However, it is difficult to allow for variable piriform fossa parameters in the articulation model, because experimental data on the activity of these regions are insufficient.

An error may also arise from the fact that we assumed the impedance of the glottis to be infinite. At the same time, it was shown that frequencies of the resonance oscillations can change when the glottis is open [5]. It is also known that 80% of women and 20% of men keep the rear part of the glottis open even when their vocal cords are closed, so that the oral cavity and subglottis region are always coupled [19]. This circumstance can also affect the resonance frequencies.

One more unaccounted source of errors is nasalization, i.e., a dropping of the soft palate and a branching of the vocal tract in the nasal region. Nasalization shifts the resonance frequencies. Its effect is however difficult to reveal by linear prediction.

The compliance of the vocal tract's walls and the presence of branchings considerably affect the resonance frequencies. The formant frequencies calculated under the assumption that the walls are perfectly rigid noticeably differ from the measured formant frequencies. Taking into account the wall compliance reduces the error in the first resonance. The parameters of the wall impedance (the loss, mass, and elasticity) can be assumed to be constant over the tract. The accuracy of calculating the first three resonance frequencies for a tract with compliant walls and piriform branchings proves to be within the accuracy of estimating the formants from the speech signal.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 03-01-0016.

REFERENCES

1. V. N. Sorokin, *Speech Synthesis* (Svyaz', Moscow, 1992).
2. V. N. Sorokin, A. S. Leonov, and A. V. Trushkin, Speech Commun. **30**, 55 (2000).
3. J. L. Flanagan, *Speech Analysis, Synthesis, and Perception* (Academic, New York, 1965; Svyaz', Moscow, 1968).
4. G. Fant, *Acoustic Theory of Speech Production* (Mouton, s'Gravenhage, 1960; Nauka, Moscow, 1964).
5. V. N. Sorokin, *Speech Formation Theory* (Radio i Svyaz', Moscow, 1985).
6. I. Titze, B. Story, and E. Hoffman, J. Acoust. Soc. Am. **100**, 537 (1996).
7. B. Story, I. Titze, and E. Hoffman, J. Acoust. Soc. Am. **104**, 471 (1998).
8. T. Baer, J. C. Gore, L. C. Gracco, and P. W. Nye, J. Acoust. Soc. Am. **90**, 799 (1991).
9. J. Dang and K. Honda, J. Acoust. Soc. Am. **101**, 456 (1997).
10. G. K. Vallabha and B. Tuller, Speech Commun. **38**, 141 (2002).
11. J. Liljencrants, DS Dissertation (Royal Inst. Tech., Stockholm, 1985).
12. B. Story, PhD Dissertation (Univ. of Iowa, 1995).
13. P. M. Morse, *Vibration and Sound* (McGraw-Hill, New York, 1936; GITTL, Moscow, 1949).
14. L. R. Rabiner and R. W. Schafer, *Digital Processing of Speach Signals* (Prentice Hall, Englewood Cliffs, N.J., 1978; Radio i Svyaz', Moscow, 1981).
15. M. A. Isakovich, *General Acoustics* (Nauka, Moscow, 1973).
16. E. Skudrzyk, *The Foundations of Acoustics. Basic Mathematics and Basic Acoustics*, 2nd ed. (Springer, New York, 1971; Mir, Moscow, 1959), Vol. 2.
17. M. M. Sondhi, J. Acoust. Soc. Am. **55**, 1070 (1974).
18. M. M. Sondhi and J. Schroeter, IEEE Trans. Acoust., Speech, Signal Process. **35**, 955 (1987).
19. D. H. Klatt, J. Acoust. Soc. Am. **67**, 971 (1980).

Translated by A. Khzmalyan

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 331–342. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 397–410.

BIOLOGICAL ACOUSTICS

Temporal Resolution and Temporal Integration of Short Pulses at the Auditory Periphery of Echolocating Animals

L. K. Rimskaya-Korsakova

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

e-mail: lkrk@akin.ru

Received October 11, 2003

Abstract—To explain the temporal integration and temporal resolution abilities revealed in echolocating animals by behavioral and electrophysiological experiments, the peripheral coding of sounds in the high-frequency auditory system of these animals is modeled. The stimuli are paired pulses similar to the echolocating signals of the animals. Their duration is comparable with or smaller than the time constants of the following processes: formation of the firing rate of the basilar membrane, formation of the receptor potentials of internal hair cells, and recovery of the excitability of spiral ganglion neurons. The models of auditory nerve fibers differ in spontaneous firing rate, response thresholds, and abilities to reproduce small variations of the stimulus level. The formation of the response to the second pulse of a pair of pulses in the multitude of synchronously excited high-frequency auditory nerve fibers may occur in only two ways. The first way defined as the stochastic mechanism implies the formation of the response to the second pulse as a result of the responses of the fibers that did not respond to the first pulse. This mechanism is based on the stochastic nature of the responses of auditory nerve fibers associated with the spontaneous firing rate. The second way, defined as the repeatition mechanism, implies the appearance of repeated responses in fibers that already responded to the first pulse but suffered a decrease in their response threshold after the first spike generation. This mechanism is based on the deterministic nature of the responses of fibers associated with refractoriness. The temporal resolution of pairs of short pulses, which, according to the data of behavioral experiments, is about 0.1–0.2 ms, is explained by the formation of the response to the second pulse through the stochastic mechanism. A complete recovery of the response to the second pulse, which, according to the data of electrophysiological studies of short-latency evoked brainstem potentials in dolphins, occurs within 5 ms, is explained by the formation of the response to the second pulse through the repetition mechanism. The time constant of temporal integration, which, according to the behavioral experiments at threshold levels of pulses, is about 0.2–0.3 ms, is explained by the integrating properties of internal hair cells, etc. It is shown that, at the high-frequency auditory periphery, the temporal integration imposes no limitations on the temporal resolution, because both integration and resolution are different characteristics of the same multiple response of synchronously excited fibers. *© 2004 MAIK "Nauka/Interperiodica".*

Temporal integration (summation) is the ability of a system to combine information in time, and temporal resolution is the ability to follow rapid temporal variations of the signal. A system with a smaller time of integration has a better temporal resolution.

Temporal integration and resolution are important characteristics of the auditory system. The auditory integration manifests itself as a decrease in the detection threshold with an increase in the stimulus duration and provides the noise immunity of the analysis of sound [1, 2]. If the threshold decrease is 3 dB when the repetition rate or duration of the stimulus is doubled, one deals with the case of energy integration. The auditory integration is of both peripheral and central origins.

The auditory temporal resolution is usually understood as the resolution of the variations of the signal envelope rather than the resolution of the fine time structure of the signal [3]. The temporal resolution may be limited at the periphery and in the central nervous system. Peripheral limitations are associated with the much investigated first stages of sound transformation, including the auditory nerve. They are determined by the inertial properties of the basilar membrane, the hair cells, and the transformations in the synapses and spiral ganglion neurons. The central limitations include the peripheral limitations and all stages of the less studied coding of sound signals at higher levels of the nervous system.

This paper considers the physiological basis of the peripheral auditory temporal resolution and temporal integration of short high-frequency pulses in echolocating animals with the aim to explain the known data obtained from behavioral and electrophysiological experiments. The manifestation of the peripheral coding properties in behavioral reactions is quite possible, because the auditory nerve is the only channel through which acoustic information is transferred to the brain. All losses that occur at the periphery are uncorrectable and, hence, may be detected in behavioral threshold experiments [4]. The relation between peripheral coding and behavioral reactions in echolocating animals

1063-7710/04/5003-0331$26.00

may be justified by the fact that a fast auditory analysis of short echo signals is necessary for survival [5]. At the same time, the principle of the functional structure of the auditory system [6] implies that the higher the carrier frequency of a stimulus or the frequency of its intensity modulation, the lower the level of the nervous system at which the coding of the stimulus terminates.

The properties of the temporal integration and temporal resolution discussed in this paper are as follows. The energy integration of short pulses in dolphins [9–14] and bats [15–17] has a time constant of 0.2–0.3 ms.[1] The measurements of the thresholds of backward temporal masking in dolphins revealed a separation of the pulsed signal and pulsed noise at intervals of 0.5 ms and a strong masking at intervals smaller than 0.3 ms [13, 18, 19]. The minimal values of the differential thresholds with respect to the interpulse interval in pairs of short pulses are obtained for dolphins at intervals of 0.05–0.5 ms [20, 21]. In discriminating paired pulses of superthreshold levels, the dolphins cease being capable of discriminating the time intervals when the latter exceed 0.2 ms [18, 19, 22, 23].

The property of the dolphin's auditory system to produce responses to repeated pulses was estimated in electrophysiological experiments on the detection of short-latency evoked brainstem potentials also called auditory brainstem responses (ABRs) [23–27]. The ABRs are formed from the multiple response of synchronously excited high-frequency auditory nerve fibers and, according to [29], reflect the properties of their combined responses. A complete recovery of the response to the second pulse was observed at an interval of 5 ms [25, 26]. An increase in the pulse intensity by 20 dB each time caused an approximately triple increase in the time of recovery, up to 20 ms. At small intervals, because of the overlapping of the ABRs arising in response to each of the pulses in a pair, the response to the second pulse was formed by the method of subtracting the response to one pulse from the response to the pair. As the interval decreased down to 0.2 ms, or even to 0.1 ms in some cases [24–27], the amplitude of the repeated response decreased but remained noticeable.

The aforementioned data of behavioral experiments were explained by their authors [9–23] from the viewpoint of the functioning of different subjective features in the auditory system, such as loudness, timbre, and pitch, which are formed by the central nervous system but are based on the properties of the frequency–time fine peripheral coding. As for the data of electrophysiological experiments, they were explained by the authors from the standpoint of the existence of a hypothetical auditory filter with a time constant of 0.2 ms [27, 28].

Is it possible to give a unified consistent interpretation of the above-mentioned data on the temporal resolution and temporal integration at the level of peripheral high-frequency coding of pairs of short pulses? Dolphins are no common laboratory animals, and studies of the responses of their real auditory nerve fibers are ruled out. However, it is possible to perform model studies. Such investigations are described in our previous publications [30–33]. Since then, some progress in this area of research has occurred and we also found new explanations for some specific features of the peripheral auditory analysis of sound signals in land animals [34–37]. All this allows us to hope to obtain the answers to the following questions. Is the analysis of the fine time structure of a short pulse possible at the periphery of the high-frequency auditory system? What properties of the peripheral coding can provide the temporal resolution of a signal within 0.2 ms [24–27] and less [13, 20, 21], which is less than the period of excitability recovery (refractoriness) of a single fiber? What stage of the peripheral processing can determine a temporal integration of about 0.2–0.3 ms, which was detected in the behavioral experiments in dolphins [9, 11, 13, 18–23] and bats [15–17]? Why does the interval value of about 0.2–0.5 ms often appear as a certain critical interval [22] in the behavioral experiments on the discrimination of paired pulses in dolphins and bats? Which stage of the peripheral processing is responsible in this case for the fact that the total time of recovery of the response to the second pulse can reach 5 ms and over, as was found from electrophysiological experiments?

The general answer to these questions is that the stimulus duration is comparable with the time constants of the basic peripheral processes of the transformation of an analog sound wave into a sequence of action potentials (spikes) of the auditory nerve fiber [5, 12, 19, 38], while the losses that occur in the course of this transformation are uncorrectable.

Let us consider a sequence of peripheral transformations of sounds. It is well known that the formation of the firing rate of auditory nerve fiber is accompanied by a series of nonlinear transformations. After a broadband filtering in the pinna and tympanum, sound is subjected to bandpass filtering in the cochlea. A linear relation between the sound pressure level and the oscillation amplitude of the basilar membrane is observed only for low sound levels. At this stage, a compression of the dynamic range is possible, and the two-tone deceleration effect takes place. A flexure of the basilar membrane causes a displacement of the cilia of external and internal hair cells, which changes the conductivity of their membranes. The resulting intracellular potential initiates the mediator secretion into the synaptic cleft between the hair cell and the dendrite of the auditory nerve fiber. Then the synaptic potential is transformed

[1] Most likely, the time constant of integration, which for dolphins reaches tens of milliseconds under the effect of long stimuli [7], takes into account both peripheral and central processes of adaptation and integration. However, the peripheral coding of short pulses is not accompanied by the auditory adaptation affecting the state of the auditory system [8].

into an intracellular potential of the auditory nerve (the generator potential) and, after the threshold transformation, into the temporal sequence of spikes (action potentials).

It is known that 90–95% of spiral ganglion cells forming the auditory nerve are connected with the internal receptor hair cells (IRHCs). In echolocating animals, a single IRHC has the maximal number of connections with the dendrites of the spiral ganglion neurons, as compared to other mammals [40, 41]. They also exhibit a spatial variation of the IRHC innervation density, which is maximal in the region of the maximal hearing sensitivity of the animal [42].

Auditory nerve fibers are nonuniform in their morphological and physiological properties and, according to their spontaneous firing rate, are separated into three groups: with low, medium, and high spontaneous firing rates [43]. Spontaneous firing rate (SFR) is the ability of the fiber to generate spikes spontaneously, in the absence of stimulus. The position of the fiber dendrite on the IRHC body (relative to the Corti channel) correlates with the size of its synaptic termination, its diameter, and the level of SFR [44]. The origin of the SFR is believed to be related to spontaneous mediator ejections from the receptor cell into the synaptic cleft. The fiber with greater SFR is characterized by a lower response threshold [45], a steeper and narrower input–output characteristic determined by the dependence of the mean firing rate on the level of the characteristic frequency tone [46], a poorer reproduction of the envelope of an amplitude-modulated signal in the response of the fiber [47, 48], and a faster recovery of excitability after the termination of the tone presentation [49]. This relation between the physiological properties of fibers and the SFR is adequately described by the model of auditory nerve fiber developed earlier in [30–32, 36, 37] and used in the present study (Fig. 1).

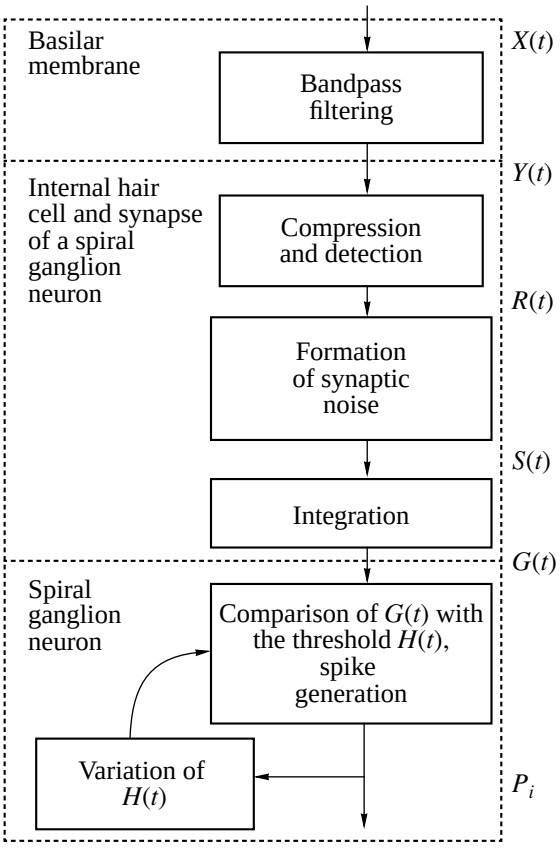


Fig. 1. Flow chart of the model of auditory nerve fiber.

The fiber model transforms a sound signal into a sequence of spikes. In the model, a sequential formation of the pulsed response of the basilar membrane, the receptor and synaptic potential of IRHCs, and the firing-rate response (a sequence of spikes) of the spiral ganglion neurons takes place. At the stage of the formation of the pulsed response of the basilar membrane $Y(t)$, a linear convolution of the input signal $x(t)$ with the pulsed characteristic of the basilar membrane filter, $h(t)$, is performed. The nonlinear transformation of the signal at the mechanoreception stage was realized by a sigmoid function reproducing the dynamics of the variation of the receptor potentials of IRHCs. The function $R(t)$ had the form

$$R(t) = R_{\max}\left[\frac{2}{1 + \exp[(\mathrm{Dis} - Y(t))/Sl]} - 1\right],$$

where $R_{\max}$ is the maximal value of the synaptic potential and Dis is its displacement, which is always equal to 0.05. It was assumed that the mechanism of the formation of the synaptic potential from the receptor potential is closely related to the mechanism responsible for the appearance of the SFR. Therefore, the slope of the sigmoid function (parameter Sl) depended on the SFR level approximately in the same way as the slope of the input–output characteristic of the real fiber did [46].

Synaptic noise $S(t)$ was formed as a random process with a mean value and a variance, each of them being proportional to the instantaneous quantity $R(t)$ subjected to detection. To reproduce the SFR, the rms deviation of the random process was increased by ξ. If the signal $X(t)$ was absent, the synaptic noise $S(t)$ had a zero mean and an rms deviation proportional to ξ. The formation of the synaptic potential $G(t)$ was terminated by the integration of $S(t)$. The integration took into account the inertial properties of the signal transformation at the mechanoreception level [50, 51] and at the membrane of the spiral ganglion neuron [52]. The time constant was 0.2 ms.

At the stage of the transformation of the synaptic potential into a sequence of spikes P_i, a comparison of $G(t)$ with the time-varying threshold $H(t)$ was performed. If $G(t)$ exceeded the threshold, the neuron model generated a spike, which caused a temporal increase in the threshold. Within an interval of time equal to the sum of the periods of absolute and relative refractorinesses, the threshold returned to its original

Parameters of auditory nerve fibers with a characteristic frequency of 70 kHz and with different spontaneous firing rates.

Type of fiber model	SFR (pulse/s)	ξ	R_{max}	Sl	τ_f (ms)	H_{of}/H_o	τ_s (ms)	H_{os}/H_o
(h) High SFR	30	0.020	0.065	0.025	2	0.4	20	0.15
(m) Medium SFR	5	0.007	0.09	0.25	3	0.5	30	0.2
(l) Low SFR	0	0.0055	0.17	1.2	4	0.7	40	0.3

level H_0. At the stage of relative refractoriness, the threshold had one fast and one slow component [53]. Each of the components of the threshold could increase by values preceding the spike generation. The threshold variation with time, $H(t)$, after the spike generation has the form

$$H(t) = H_0 + H_f \exp\left[-\frac{t-\tau_a}{\tau_f} + \Omega \mathrm{Sign}(\tau_a - t)\right] + H_s \exp\left[-\frac{t-\tau_a}{\tau_s} + \Omega \mathrm{Sign}(\tau_a - t)\right], \quad (1)$$

where $H_f = H_{of} + dH_f$; $H_s = H_{os} + dH_s$; H_{of}; H_{os};H_f, and H_s are constant and variable values of the fast and slow components of the neuron threshold at the stage of relative refractoriness; dH_f and dH_s are the values of the threshold components at the instant preceding the spike generation; τ_a is the absolute refractoriness period always equal to 0.5 ms; τ_f and τ_s are the time constants of the two threshold components during the period of relative refractoriness; Ω is a constant that determines the threshold at the stage of absolute refractoriness; and $\mathrm{Sign}(t) = 1$ for $t < 0$ and $\mathrm{Sign}(t) = 0$ for $t > 0$. The threshold function $H(t)$ allows one to preset the refractoriness and adaptation properties in the fiber models, where the refractoriness is the change in the excitability of the fiber after the spike generation and the adaptation is the decrease in the firing rate during the stimulus presentation.

The central frequency of the filter of the basilar membrane was equal to 70 kHz and determined the characteristic frequency of the fiber model. The known relation between the SFR level, the steepness of the input–output characteristic, and the fiber ability to reproduce the modulation of sound [36, 37] was achieved by satisfying the following conditions: the smaller the parameters ξ, τ_f, τ_s, H_{of}/H_o, H_{os}/H_o, and ξ determining the SFR level, the greater the parameter Sl determining the slope of the function $R(t)$. The parameter R_{max} was chosen so that the mean firing rate in the saturation regime did not exceed 400 pulse/s. The value of the threshold H_0 was always equal to 0.02. Therefore, the scatter in the response thresholds in the models with different SFRs occurred because of the difference in the SFR and the steepness of the function $R(t)$. Other values of the parameters of high-frequency fiber models with different SFRs are given in the table. The stimuli were pairs of pulses whose duration was 35 μs and the central frequency of the spectrum, 70 kHz. The interval between pulses in a pair was varied within 0.05–25 ms. The time step in the calculations was 1 μs.

The SFR is an important parameter, depending on which the models of auditory nerve fibers, as well as real fibers, exhibit different response properties.

Figure 2a presents the input–output characteristics, i.e., the dependence of the mean firing rate on the tone level, for the model of auditory nerve fiber with a characteristic frequency of 70 kHz. These characteristics were obtained in response to a tone of the same frequency with a duration of 40 ms. For the models of high-frequency fibers, as for the models of low-frequency fibers [36, 37], the input–output characteristic is steeper and the dynamic range is narrower when the SFR level is higher. Since, at the threshold levels of stimuli, only the fibers with low thresholds and high SFR are involved in the response,[2] their threshold was taken to be equal to 0 dB.

In the high-frequency fiber model, a short pulse with a duration of 35 μs may cause (due to refractoriness) only one spike. Therefore, the input–output characteristic (Fig. 2b) obtained in response to a short pulse was estimated by the dependence of not the mean firing rate but the firing rate probability, i.e., the ratio of the number of spikes to the number of stimulus presentations, on the pulse level.

The period of paired pulse presentation was large to avoid the effect of adaptation and refractoriness on the response of the fiber to each subsequent pair of pulses. Therefore, the probability of spike generation in one fiber in response to repeated stimuli was equal to the probability of spike generation in multiple identical fibers in response to one stimulus presentation.

It was found that the steepness of the second input–output characteristic of the fiber (Fig. 2b) was greater when the SFR level was lower. Although the response thresholds of the fiber models under the action of short pulses increased by almost 20 dB, the lowest response

[2] The inertial property of one fiber is compensated by the response of many fibers of the auditory nerve with a probabilistic (due to the SFR) nature of responses. The lowest thresholds correspond to fibers with high SFR. As will be seen below, a high SFR provides the readiness of approximately the same number of fibers for spike generation at any instant of time.

threshold was retained by the fiber model with a high SFR.

The modeling allows one to study the independent effects of different processes on the steepness of the input–output characteristic, namely, such processes as spontaneous spike generation, formation of receptor potentials of a hair cell $R(t)$, and time variation of the spike generation threshold $H(t)$ due to the adaptation and refractoriness. The difference in the behavior of the steepness of the input–output characteristic as a function of the SFR level under the effect of stimuli of different durations is caused by the fact that, under the effect of a tone, the steepness is determined by the dependence of the steepness of the receptor potential on the tone level (parameter SL of the function $R(t)$) (Fig. 2a), while under the effect of a short pulse, the steepness of the input–output characteristic depends on the SFR (parameter ξ) (Fig. 2b).

Evidently, in the multitude of auditory nerve fibers, a response to the second pulse of a pair can be formed in only two ways. One way, which can be called the stochastic mechanism, includes the formation of the response to the second pulse from the responses of the fibers that did not respond to the first pulse. This mechanism is based on the probabilistic (stochastic) nature of the SFR-related responses of multiple auditory nerve fibers. The second way, which can be called the repetition mechanism, includes the appearance of repeated responses in the fibers that already produced a response and that reduced their response threshold due to the spike generation under the effect of the first pulse. This mechanism is based on the refractoriness-related deterministic nature of the responses.

Figure 3 shows the distinctions in the formation of responses to the second pulse of a pair for the intervals of 0.2 and 2 ms in the fiber models with different SFRs. In both cases, the pulsed responses of the basilar membrane model, $Y(t)$, are completely separated in time. Because of the integration, the receptor $R(t)$ and synaptic $G(t)$ potentials are smoothed out. When intervals between pulses in pairs are smaller than 0.4–0.5 ms, these potentials interact. Therefore, the amplitude of the receptor (synaptic) potential arising in response to the second pulse of a pair is greater than that corresponding to the first pulse (Fig. 3a). An increase in the pulse level increases the time of interaction of the receptor potentials. At the same time, the transformation of the synaptic potential into a sequence of spikes in a multitude of identical fibers is accompanied by a decrease in the duration of the pulsed response of the fiber, so that the spike generation is associated with only the rising segment of the synaptic potential $G(t)$ [30–32].

In response to a single presentation of a pair of pulses with an interval of 0.2 ms (Fig. 3a), a single fiber may generate no more than one spike (because of the refractoriness) at an arbitrary instant of time (because of the SFR). However, when the interval is 2 ms

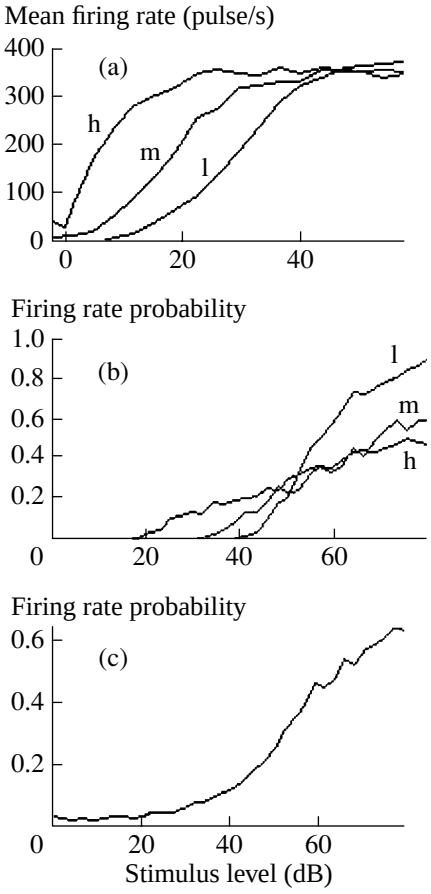


Fig. 2. Input–output characteristics of the models of auditory nerve fibers with high (h), medium (m), and low (l) spontaneous firing rates and a characteristic frequency of 70 kHz. (a) Dependence of the mean firing rate in a steady-state regime on the level of a 70-kHz tone signal with a duration of 40 ms. (b) Dependence of the firing rate probability due to the presentation of a short pulse with a duration of 35 μs and a spectrum with a central frequency of 70 kHz on the pulse level. (c) Total firing rate probability in three groups of fiber models with different SFRs.

(Fig. 3b), more than one spike may be generated, because, after the generation of the first spike, the synaptic potential arising in response to the second pulse may exceed the now reduced response threshold of the fiber.

At an interpulse interval of 0.2 ms, the response to the second pulse is formed only by the stochastic mechanism. The total single response of a multitude of fibers reproduces not only the fine time structure of a pair of pulses (the characteristic frequency of the fiber model is extracted) but also its envelope (the receptor-enhanced response to the second pulse of the pair is extracted) (Fig. 3a, plots h, m, l, t). When the pulse levels are low, the probabilities of responses to individual pulses of a pair are proportional to the amplitudes of the corresponding receptor potentials. The lower the SFR level in the fiber model, the better the pulse response discriminates the amplitude of the second pulse from

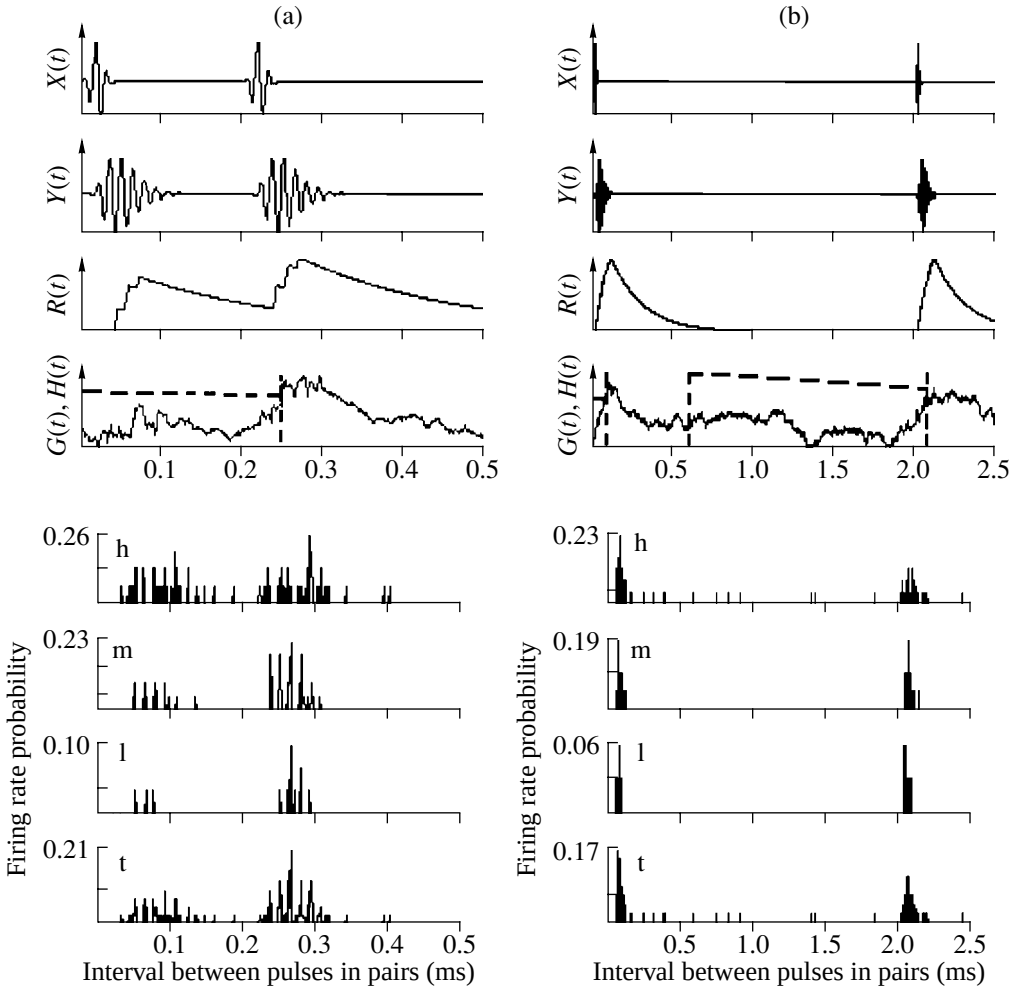


Fig. 3. Responses of different units of the auditory nerve fiber model to pairs of pulses with intervals of (a) 0.2 and (b) 2 ms. From top to bottom: $X(t)$ is the temporal profile of the stimulus, $Y(t)$ is the pulsed response of the basilar membrane model, $R(t)$ is the smoothed temporal profile of the receptor potential, $G(t)$ and $H(t)$ are the time variations of the synaptic potential (the solid line) and the threshold (the dashed line); the vertical dashed line refers to the instant of spike generation and the termination of the state of absolute refractoriness of the neuron. The lower plots show the pulsed responses of the fibers with high (h), medium (m), and low (l) spontaneous firing rates and the total response of three groups of fibers (t). The level of paired pulses is 45 dB.

the first pulse [30, 31], which corresponds to the properties of the input–output characteristic obtained for the case of a single pulse (Fig. 2b) rather than a tone stimulus (Fig. 2a).

When the interval in a pair is 2 ms, the amplitudes of the receptor potentials arising in response to each of the pulses are identical (Fig. 3a, $R(t)$). The response to the second pulse in the multitude of fibers is formed through both the stochastic and repetition mechanisms (Fig. 3b). However, in fibers with different SFRs, the contributions of these mechanisms to the formation of the response to the second pulse are different. Therefore, the response of the fiber with a high SFR and low thresholds to the second pulse of a pair (Fig. 3b, plot h) proved to be smaller than the response to the first pulse. Fibers with medium and low SFRs exhibit equal probabilities of responses to the fist and second pulses (Fig. 3b, plots m, l). The integrated response of three groups of fibers with different SFRs is determined by the properties of the group that has the majority of its fibers involved in the response at a given stimulus level (Fig. 3b, plot t). Therefore, the firing rate probabilities of all three groups of fibers proved to be closer to the firing rate probabilities of fibers with high and medium SFRs (Fig. 3b, plots h, m).

Let us try to estimate the individual contributions of the two aforementioned mechanisms to the formation of the response to the second pulse and also determine their dependences on the pulse levels, interpulse interval, and SFR level in the fiber.

Usually, the probabilities of the appearance of a response to individual pulses of a pair are estimated by the ratio of the numbers of spikes $N1$ and $N2$ arising within certain time intervals to the number of pair presentations. The number $N2$ includes the number of spikes arising through both mechanisms of the forma-

tion of the response to the second pulse. For a separate estimation of each of these mechanisms, a summation of spikes that occur in response to the second pulse was performed for the case of the absence ($Ns2$) and presence ($Nr2$) of a response to the first pulse. The dependence of the ratio $N2/N1$ on the interval between the pulses of a pair represents the recovery function. The contributions of stochastic and repetition mechanisms to the total recovery function are estimated by the dependences of the ratios $Ns2/N1$ and $Nr2/N1$ on the interpulse interval, respectively. All three recovery functions are shown in Fig. 4.

Model experiments show that the ratio $Ns2/N1$ does not depend on the interval but depends on the level of the receptor potentials alone (Fig. 4b). In low-threshold fibers with a high SFR, the probability of spike generation per one pulse does not exceed 0.5 or 0.6 for any pulse levels (see Fig. 2b). This provides good conditions for the formation of the response to the second pulse for any pulse levels and any interpulse intervals (see Fig. 4b, plots h, m). The role of the stochastic mechanism decreases as the level of SFR is lowered (Fig. 4b, plot l). An increase in the pulse level increases the probability of the response to the first pulse and reduces that for the second pulse. At high pulse levels, the response to the second pulse will never recover if only the stochastic mechanism provides the formation of the response to the second pulse.

The repetition mechanism depends on the interpulse interval, the level of the paired pulses, and the SFR of the fiber (Fig. 4b). The ratio $Nr2/N1$ is greater for higher pulse levels, greater interpulse intervals, and the lower SFRs.

A combined effect of the two mechanisms providing the formation of the response to the second pulse is shown in Fig. 4a. As follows from Fig. 4b, an increase in the pulse level reduces the effect of the stochastic mechanism and enhances the effect of the repetition mechanism. In this case, the effect of the repetition mechanism is stronger when the pulse level is lower and the SFR of the fiber is lower. Fibers with high and medium SFRs exhibit a dependence of the recovery time on the pulse level (Fig. 4a, plots h, m). In fibers without SFR (Fig. 4a, plot l), such a dependence is absent. When the pulse level reaches 80 dB (Fig. 4a, plot l), the stochastic mechanism proves to be almost completely displaced by the repetition mechanism (according to Fig. 2b, plot l, at a level of 80 dB, the probability of the response to one pulse is 0.8–0.9). Therefore (Fig. 4a, plot l), at a pulse level of 80 dB (at which only the repetition mechanism is working), the steepness of the recovery function is greater than at a level of 60 dB (at which both mechanisms are active).

The combined action of the two mechanisms at relatively low paired pulse levels may result in the response to the second pulse exceeding the response to the first pulse (see Fig. 4a, plots h, m: the values of the recovery function at intervals greater than 2 ms proved to be greater than unity).

At a given pulse level, the total response of three groups of fibers with different SFRs (Fig. 4a, plot t) depends on the properties of the group that has the majority of its fibers involved in the response (as was mentioned above). Therefore, at low levels, the total recovery function of fibers (Fig. 4a, plot t) reproduces the properties of the firing-rate response for the fibers with a high SFR (Fig. 4a, plot h), and at higher levels, for the fibers without SFR (Fig. 4a, plot l). In the total response of three groups of fibers, the recovery time of the response to the second pulse at a stimulus level of 40 dB is about 3 ms, at a stimulus level of 60 dB, 6.5 ms, and at a level of 80 dB, 20 ms.

It should be pointed out that, at intervals of 0.2–0.5 ms (smaller than the period of absolute refractoriness), the response to the second pulse is formed through the stochastic mechanism alone. Therefore, the higher the level of stimuli, the smaller the values of the recovery functions are. However, at threshold levels and low levels of the pulses (no higher than 20 dB), the responses of the fiber models with high and medium SFRs, as well as the total response of three groups of fibers, may exhibit a receptor amplification of the second pulse (Fig. 4a, plots h, m, t, and Fig. 5 described below).

Since the time integration in behavioral experiments was estimated on the basis of the threshold experiments, let us consider the responses of low-threshold fibers with high SFR under the effect of paired pulses also with threshold levels (Fig. 5). At the lowest pulse level, owing to the interaction of the receptor potentials of hair cells, the variation of the interpulse interval is accompanied by the following variation of the total number of spikes singly generated in the multitude of fibers in response to a pair of pulses: the number of spikes first smoothly increases, then, at intervals of about 0.1–0.15 ms, it passes through a maximum, and then, as the interval increases further, it decreases. The dependence of the number of singly generated spikes on the interval value exhibits an integration corresponding to a receptor enhancement of the second pulse. Such a behavior of the aforementioned characteristic is related to the ratio of the numbers of single and repeated spikes arising in the multitude of excited fibers. Therefore, an increase in the pulse level (see Fig. 5) leads to a decrease in the interval values at which the integration effect is observed. A similar situation occurs when, instead of pairs of short pulses, long stimuli are presented (not shown in the figures).

The model experiments demonstrate (Fig. 3) how the peripheral processing analyzes the fine time structure of a pair of pulses (the characteristic frequency of the fiber models is extracted) and their envelope (the envelope of a pair of pulses is extracted). However, a short high-frequency pulse has a broad spectrum, which simultaneously excites a multitude of fibers with

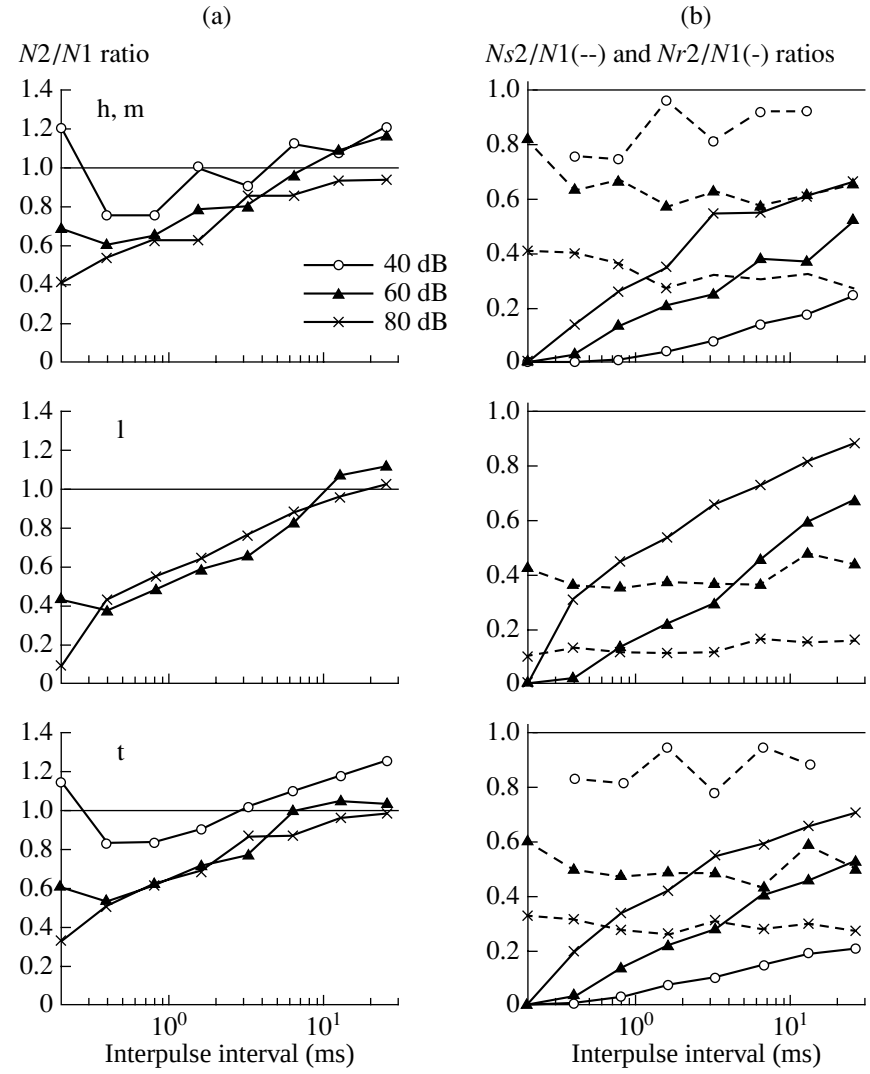


Fig. 4. Recovery functions of the response to the second pulse for the fiber models with high and medium (h, m) spontaneous firing rates and low (l) spontaneous firing rate versus the interpulse interval in pairs. Plot t shows the total recovery functions obtained for three groups of fibers. The ordinate axes represent (a) the ratio $N2/N1$, where $N1$ and $N2$ are the numbers of spikes arising in response to the first and second pulses of the pair, and (b) the ratios $Ns2/N1$ (the dashed lines) and $Nr2/N1$ (the solid lines), where $Ns2$ and $Nr2$ are the numbers of spikes arising in response to the second pulse in the absence and presence of response to the first pulse, respectively. The pulse level serves as a parameter, the values of which are indicated in the plot. The horizontal line corresponds to the 100% recovery of the response to the second pulse.

different characteristic frequencies. Therefore, we do not discuss the question of how the auditory system uses the information on the fine time structure of the signal that is extracted by the fibers with different characteristic frequencies. Here, we only discuss the properties of the peripheral temporal analysis of the envelope of a pair of pulses. These properties are necessarily retained in the properties of the response of the multitude of fibers with different characteristic frequencies, because, in the multitude of channels, a short pulse evokes a synchronized firing rate of all high-frequency fibers [29] and because the maximal innervation of IRHCs by the auditory nerve fibers falls within a limited range of their characteristic frequencies near the maximal sensitivity of hearing of the animal [42].

First, let us compare the responses of the fiber models with the responses of real auditory nerve fibers that are known from the literature.

In the model experiment, we calculated two input–output characteristics for the cases of presentation of long tone signals and short pulses (Figs. 2a, 2b). The steepness of the profile of the first characteristic (Fig. 2a) is independent of the characteristic frequency of the fiber model. Our previous studies showed that the input–output characteristics obtained for fiber models with different SFRs adequately describe not only the real dependences of the mean firing rate of auditory

nerve fibers on the tone level but also the abilities of the fibers to reproduce amplitude modulation [36, 37].

The second input–output characteristic (Fig. 2b) reproduces the dependence of the probability of the response of the fiber model on the level of a short pulse in the case of its multiple presentation and corresponds to the probability of a response appearance in the multitude of singly excited fibers. An integrated input–output characteristic for three groups of fiber models (Fig. 2c) can be qualitatively compared with the dependence of the ABR amplitude on the level of a short pulse [25–27], because the short-latency evoked brainstem potentials (i.e., ABRs) in dolphins are formed as a result of synchronous responses of many high-frequency auditory nerve fibers to a short high-frequency pulse. Since such a pulse can evoke only one spike in every high-frequency fiber (because of the refractoriness), the ABR amplitude is proportional to the number of single spikes in the multitude of excited fibers.

The dependence of the ABR amplitude on the pulse level was approximated by a straight line [26], and the dynamic range determined from it was 50 dB. If the stimulus was a long noise pulse instead of a short pulse, the level at which the minimal ABR amplitude was observed proved to be 20 dB lower. For the fiber model at hand, the integrated input–output characteristic (Fig. 2c) can also be approximated by a straight line, but starting from the level of 30 dB. In this case, the dynamic range also reaches 50 dB. The response threshold of the low-threshold fiber model with a high SH was found to be 20 dB lower in the case of the presentation of a long tone signal (Fig. 2a) compared to the response threshold in the case of the presentation of a short pulse (Fig. 2b). The 20-dB difference is not only a consequence of the energy integration but also appears as a result of the summation of probabilities of responses, because an increase in the stimulus duration increases the probability of its detection [54, 55].

The formation of the response to the second pulse by the stochastic mechanism in real auditory nerve fibers is shown in Fig. 1 of paper [56]. The authors obtained post-stimulus histograms of the responses to paired bipolar electric pulses with an interpulse interval of 0.1 ms and with different pulse levels. This interval was smaller than the period of absolute refractoriness. At low pulse levels, the fiber responded only to the second pulse, which was amplified by the residual depolarization of the dendrite membrane under the effect of the first pulse. At medium levels, a response to each of the pulses in the pair was observed. At higher levels, the response to the first pulse was present while the response to the second pulse disappeared. In our model, the response to the second pulse is formed in the same way through the stochastic mechanism (Fig. 3) [24, 20, 31, 32].

The properties of the recovery functions of auditory nerve fibers with different SFRs were studied in the cat by presenting short high-frequency paired pulses with

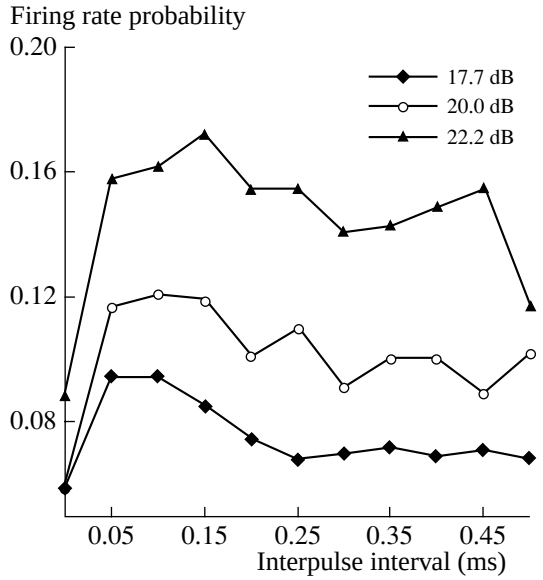


Fig. 5. Dependence of the firing rate probability for a fiber model with a high spontaneous firing rate on the interpulse interval in the case of the presentation of a threshold-level pair of pulses (the levels are indicated as parameters).

different pulse levels and different intervals from 1 to 32 ms [57, 58]. The responses of the fibers with characteristic frequencies of 0.5–60 kHz were observed.

Real and model fibers showed identical dependences of the duration of the pulsed response on the pulse level. For real fibers with SFR [57], the recovery time of the response to the second pulse was greater when the pulse level was higher. For fibers without SFR, the recovery time was almost independent of the pulse level. At high pulse levels, the recovery time was greater in fibers with high SFR compared to fibers without SFR. The response recovery period was within 3–20 ms [57] and did not coincide with the recovery period of the response thresholds of auditory nerve fibers in the experiments on the masking of tone pulses [49], where this period was equal to 150–200 ms. The difference between the recovery periods of responses to short [57] and long [49] stimuli was explained in [57] by the fact that the process of recovery of the response to the second stimulus reflected the refractoriness properties of the fibers in the first case and the adaptation properties in the second case.

In model fibers, as in real ones, the recovery period varied within 3–20 ms and was related to the refractoriness properties. The model also adequately described the dependence of the recovery function of the response to the second pulse on the SFR level, which was observed for real fibers [56, 57]. Therefore, the results reported in [57] can be interpreted from the viewpoint of the existence of the stochastic and repetition mechanisms underlying the formation of the response to the second pulse. These mechanisms are associated with the stochastic and deterministic components of the

responses of auditory nerve fibers. The model experiment showed that the higher the SFR in the fiber the better the conditions are for the formation of the response to the second pulse through the stochastic mechanism (Fig. 4b). However, this mechanism provides no recovery (if it acts alone) or prolongs the recovery period (if it acts together with the repetition mechanism) of the response to the second pulse in the case of high pulse levels. For example, the recovery period in fibers with high SFR was greater than in fibers without SFR (see Fig. 4a, plots h, m and Fig. 4a, plot l at pulse levels of 80 dB). The recovery period in fibers without SA was independent of the pulse level (Fig. 4a, plot l), because in these fibers the repetition mechanism displaced the stochastic mechanism at high pulse levels.

The behavior of the integrated recovery functions obtained by modeling (Fig. 4a, plot t) is very similar to the behavior of the recovery functions obtained in recording the integrated action potentials of the auditory nerve fibers in bats [58]: the recovery of the amplitude of the repeated response occurred within 20 ms at high pulse levels and within shorter periods at lower pulse levels.

The behavior of the same model integrated recovery functions (Fig. 4a, plot t) is similar to the behavior of the recovery functions obtained from the measurements of the ABRs in dolphins in response to pairs of short acoustic pulses [25–27]. The experimentally measured ABRs exhibit a dependence of the recovery period of the response to the second pulse on the level of the paired pulses. The variation of this period was estimated in the model experiment to be within 3 to 20 ms when the pulse level varied by 40 dB, and these values agree well with real experimental data [25–27]. The increase in the recovery period with increasing pulse level is caused by the change in the ratio between the numbers of fibers that are involved in the response and have different SFR levels, and hence, different physiological properties.

If the stimuli were pairs of pulses with intervals of 0.2–0.4 ms, the ABR amplitude arising in response to the second pulse decreased with increasing pulse levels [27]. As follows from the model experiment and from the data reported in [56], such a decrease occurs when the response to the second pulse is formed by the stochastic mechanism.

In [57, 59], it was shown that, when the levels of pulses are low and the intervals are about 20 ms or greater, the response to the second pulse may exceed the response to the first pulse. The model experiment shows that such an excess appears if the response to the second pulse is formed by the stochastic and repetition mechanisms simultaneously. An enhancement of the response to the second pulse is favored by the integration of the statistical and deterministic components of the responses of many auditory nerve fibers. Such an enhancement is not a manifestation of energy integration but is a consequence of the statistical nature of spike generation in neurons and neuron ensembles.

An example of the temporal energy integration is shown in Fig. 5. In the model experiment at the threshold levels of pulses, the integration manifests itself as an increase in the total number of spikes generated in response to a pair of pulses. The integration proved to be a consequence of the integrating properties of IRHCs, because in the IRHC model, according to the data of [50, 51], the integration time constant was 0.2 ms. The result shown in Fig. 5 can be qualitatively compared with the dependence of the ABR amplitude on the acoustic pulse duration that was obtained for dolphins [24, 25]. In dolphins, a noticeable temporal integration was observed at the lowest levels of stimuli with durations no greater than 0.3 ms. With such durations, the stimuli still had the form of short pulses.

To compare the properties of the aforementioned energy integration (Fig. 5) with the data of behavioral experiments, an additional model calculation was performed. In the fiber model, the detection threshold of a single pulse was assumed to be the level at which the total number of spikes arising in response to the signal was more than three times greater than the average number of spikes corresponding to the spontaneous firing rate. Under this assumption, the detection threshold of a pair of pulses with an interval of 0.1 ms was 3–4 dB lower than that of a single pulse, and with an interval of 0.2–0.3 ms, it was 1 dB lower. This result agrees well with the results of behavioral experiments on dolphins [11, 12] and bats [15–17], in which approximately the same kind of integration was observed at threshold levels with the same intervals in pairs of pulses. According to [17], this integration is determined by the integration of pulsed responses on the basilar membrane, but our model experiments show that this integration is caused by the integrating properties of the IRHCs.

Thus, in the model experiments, it was found that, depending on the pulse level and interpulse interval in pairs, the formation of the response to the second pulse may occur with the participation of the stochastic mechanism, or both stochastic and repetition mechanisms, or, alternatively, the repetition mechanism alone.

In the case of low or medium pulse levels and intervals smaller than the period of absolute refractoriness, only the stochastic mechanism is working (at high pulse levels, it does not work). This mechanism is based on the statistical nature of the responses of multiple auditory nerve fibers and the properties of the analog-to-pulse transformation, which shorten the IRHC response prolonged by the temporal integration. The single response produced by multiple auditory nerve fibers provides a good temporal resolution of pairs of short pulses.

In the case of different pulse levels, the period of a complete recovery of the response to the second pulse varies from 3 to 20 ms and is determined by the combined action of the stochastic and repetition mecha-

nisms. At high pulse levels, the repetition mechanism completely displaces the stochastic mechanism in the formation of the response to the second pulse.

Now, let us use these conclusions to interpret the results of the cited behavioral experiments on the temporal resolution of pairs of short pulses. The interval value of 0.2–0.5 ms, which is often encountered in behavioral experiments as a certain critical interval, is a consequence of the high-frequency auditory analysis of stimuli whose duration is comparable with the time constants of various transient processes occurring at the periphery. The model takes into account some of these processes, for example, the processes that occur on the basilar membrane, in the IRHCs, and in the spiral ganglion neurons. The oscillations of the correct discrimination of pairs of short pulses with intervals smaller than 0.1–0.15 ms, which were observed in behavioral experiments with dolphins [14, 23], may be a consequence of the interaction of the pulsed responses of the basilar membrane to each of the pulses in a pair [31]. The presence of the absolute refractoriness period of 0.5 ms determines a certain boundary within which the stochastic mechanism provides the formation of the response to the second pulse, while outside this boundary, both stochastic and repetition mechanisms are working. It seems to be quite logical that an abrupt transition from one mechanism to another in the formation of the response to the second pulse of a pair may cause an abrupt change in the behavioral reactions, such as, for example, the termination of the discrimination of time intervals exceeding 0.2 ms in dolphins [18, 19, 22, 23]. The total recovery time of the response to the second pulse, which was determined in electrophysiological experiments to be 5 ms and over [25, 26], is explained by the recovery and excitability properties of each of the multiple synchronously excited high-frequency fibers or by the properties of the repetition mechanism of the formation of response to the second pulse. The temporal integration of about 0.2–0.3 ms, which was revealed in the behavioral experiments with dolphins and bats, can be explained by the integrating properties of the IRHCs.

It should be noted that the peripheral mechanisms that determine the temporal integration do not limit the peripheral temporal resolution (Figs. 3, 5), because each of these mechanisms is based on the same multiple response of synchronously excited high-frequency fibers.

In conclusion, it should be noted that this study does not seek a relation between the parameters of the firing rate of auditory nerve fibers and the subjective qualities of sound or the properties of hypothetical auditory filters, because none of them are formed in the auditory nerve. However, the manifestation of the properties of peripheral coding in the behavioral reactions of animals, as was noted above, is possible when the stimuli are chosen so as to estimate the threshold properties of the auditory system. In many cases, pairs of pulses are such stimuli for echolocating animals, because the durations of these stimuli are comparable with the durations of peripheral transient processes. The estimates obtained for the temporal resolution of pairs of pulses from different behavioral and electrophysiological experiments prove to be different, because the chosen parameters of stimulation are critical with respect to different peripheral processes.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project nos. 00-04-49311 and 03-04-48746.

REFERENCES

1. R. Plomp and M. A. Bouman, J. Acoust. Soc. Am. **31**, 749 (1959).
2. J. Zwislocki, J. Acoust. Soc. Am. **32**, 1046 (1960).
3. N. F. Viemeister and C. J. Plack, in *Springer Handbook of Auditory Research Human Psyhophysics*, Ed. by W. Yost, A. Popper, and R. Fay (Springer, Berlin, 1993), pp. 116–153.
4. N. G. Bibikov, *Description of Sound Features by Neurons of the Auditory System of Land Vertebrates* (Nauka, Moscow, 1987).
5. É. Sh. Aĭrapet'yants and A. I. Konstantinov, *Echolocation in Nature* (Nauka, Leningrad, 1974).
6. A. G. Vasil'ev, *Echolocating Auditory System of Bats* (Leningr. Gos. Univ., Leningrad, 1983).
7. C. Johnson, J. Acoust. Soc. Am. **43**, 757 (1967).
8. J. Eggermont, Hear. Res. **157**, 1 (2001).
9. A. V. Zanin, G. L. Zaslavskiĭ, and A. A. Titov, in *Proceedings of 9th All-Union Acoustical Conference* (1977), Sect. Ts, p. 21.
10. P. W. B. Moore, R. W. Hall, W. A. Friedl, and P. E. Nachtigall, J. Acoust. Soc. Am. **76**, 314 (1984).
11. W. W. L. Au, P. W. B. Moore, and D. A. Pawloski, J. Acoust. Soc. Am. **83**, 662 (1988).
12. W. W. L. Au, *The Sonar of Dolphin* (Springer, New York, 1993).
13. N. A. Dubrovskiy, in *Sensory Abilities of Cetaceans: Laboratory and Field Evidence*, Ed. by J. A. Thomas and R. A. Kastelein (Plenum, New York, 1990), pp. 233–254.
14. N. A. Dubrovsky, in *Black Sea Aphaline*, Ed. by V. E. Sokolov and E. V. Romanenko (Nauka, Moscow, 1997), pp. 544–574.
15. J. A. Simmons, E. G. Freedman, S. B. Stevenson, *et al.*, J. Acoust. Soc. Am. **86**, 1318 (1989).
16. L. Wiegrebe and S. Schmidt, Hear. Res. **102**, 35 (1996).
17. P. Weissenbacher, L. Wiegrebe, and M. Kossl, J. Comp. Physiol. A **188**, 147 (2002).
18. V. A. Vel'min and N. A. Dubrovsky, Dokl. Akad. Nauk SSSR **225**, 229 (1975).
19. V. M. Bel'kovich and N. A. Dubrovsky, *Sensory Basis of Orientation in Cetaceans* (Nauka, Leningrad, 1976).
20. G. L. Zaslavskiĭ and V. A. Ryabov, in *Proceedings of 9th All-Union Acoustical Conference* (1977), Sec. Ts, p. 13.

21. G. L. Zaslavskiĭ and V. A. Ryabov, in *Proceedings of 11th All-Union Acoustical Conference* (1991), Sect. Ts, p. 33.
22. V. A. Vel'min and N. A. Dubrovsky, Akust. Zh. **22**, 522 (1976).
23. N. A. Dubrovsku, P. S. Krasnov, and A. A. Titov, in *Abstracts of Papers of VII All-Union Workshop on Marine Mammals* (Simferopol, 1978), p. 114.
24. N. G. Bibikov, L. K. Rimskaya-Korsakova, A. V. Zanin, and N. A. Dubrovsky, in *Electrophysiology of the Sensory Systems of Marine Mammals*, Ed. by V. E. Sokolov (Nauka, Moscow, 1986), pp. 56–84.
25. V. V. Popov and A. Ya. Supin, in *Electrophysiology of the Sensory Systems Marine Mammals*, Ed. by V. E. Sokolov (Nauka, Moscow, 1986), pp. 85–106.
26. V. V. Popov and A. Ya. Supin, J. Comp. Physiol. A **166**, 385 (1990).
27. A. Ya. Supin and V. V. Popov, J. Acoust. Soc. Am. **97**, 2586 (1995).
28. V. V. Popov and A. Ya. Supin, J. Acoust. Soc. Am. **102**, 1169 (1997).
29. M. Don and J. Eggermont, J. Acoust. Soc. Am. **63**, 1084 (1978).
30. L. K. Rimskaya-Korsakova, Akust. Zh. **35**, 887 (1989) [Sov. Phys. Acoust. **35**, 516 (1989)].
31. L. K. Rimskaya-Korsakova and N. A. Dubrovsky, Sens. Sist. **4**, 92 (1990).
32. N. A. Dubrovsky and L. K. Rimskaya-Korsakova, in *Marine Mammal Sensory Systems*, Ed. by J. Tomas, R. Kastelein, and A. Supin (Plenum, New York, 1992), pp. 223–233.
33. L. K. Rimskaya-Korsakova, Sens. Sist. **12**, 376 (1998).
34. N. A. Dubrovskiĭ and L. K. Rimskaya-Korsakova, Akust. Zh. **43**, 492 (1997) [Acoust. Phys. **43**, 421 (1997)].
35. N. A. Dubrovskiĭ and L. K. Rimskaya-Korsakova, Akust. Zh. **44**, 213 (1998) [Acoust. Phys. **44**, 173 (1998)].
36. L. K. Rimskaya-Korsakova, V. N. Telepnev, and N. A. Dubrovskiĭ, Ross. Fiziol. Zh. im I.M. Sechenova **89**, 700 (2003).
37. L. K. Rimskaya-Korsakova, Akust. Zh. **50** (2) (2004) (in press).
38. *Auditory System*, Ed. by Ya. A. Al'tman (Nauka, Leningrad, 1990), pp. 14–42.
39. N. G. Bibikov, in *Mechanisms of Functioning of the Peripheral Elements of Auditory Pathway*, Ed. by M. A. Ostrovskiĭ (VINITI, Moscow, 1988), Itogi Nauki Tekh., Vol. 39, pp. 122–211.
40. L. S. Bogoslovskaya and G. N. Solntseva, *Auditory System of Mammals: a Comparative Physiological Sketch* (Nauka, Moscow, 1979).
41. A. Berglund and D. Ryugo, J. Comp. Physiol. A **255**, 560 (1987).
42. E. Keithley and R. C. Schreiber, J. Acoust. Soc. Am. **81**, 1036 (1987).
43. M. C. Liberman, J. Acoust. Soc. Am. **63**, 442 (1978).
44. M. C. Liberman and M. E. Oliver, J. Comp. Neurol. **223**, 163 (1984).
45. C. D. Geisler, L. Deng, and S. Greenberg, J. Acoust. Soc. Am. **77**, 1102 (1985).
46. I. M. Winter, D. Robertson, and G. K. Yates, Hear. Res. **45**, 191 (1990).
47. W. S. Rhode and S. Greenberg, J. Neurophysiol. **71**, 1797 (1994).
48. R. D. Frisina, Hear. Res. **158**, 1 (2001).
49. E. M. Relkin and J. R. Doucet, Hear. Res. **55**, 215 (1991).
50. R. Patuzzi and P. M. Sellick, J. Acoust. Soc. Am. **74**, 1734 (1983).
51. I. J. Russel and P. M. Sellick, J. Physiol. (London) **284**, 261 (1978).
52. K. G. Hill, J. Comp. Physiol. A **152**, 475 (1982).
53. N. G. Bibikov and G. A. Ivanitskiĭ, Biofizika **30**, 141 (1985).
54. N. F. Viemeister and G. H. Wakefield, J. Acoust. Soc. Am. **90**, 858 (1991).
55. J. Tougaard, J. Comp. Physiol. A **183**, 563 (1998).
56. C. van den Honert and P. H. Stypulkowski, Hear. Res. **29**, 207 (1987).
57. K. Parham, H. B. Zhao, and D. O. Kim, J. Neurophysiol. **76**, 17 (1996).
58. K. Parham, H. B. Zhao, and D. O. Kim, Hear. Res. **125**, 131 (1998).
59. J. H. Siegel and E. M. Relkin, Hear. Res. **29**, 169 (1987).

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 343–349. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 411–418.

BIOLOGICAL ACOUSTICS

Some Results of Studying the Acoustics of Dolphins

E. V. Romanenko

Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninskiĭ pr. 33, Moscow, 119071 Russia

e-mail: sevin@orc.ru

Received November 25, 2003

Abstract—An experimental study of the echolocation ability of dolphins (*Tursiops truncatus*) is performed in the presence of correlated and uncorrelated broadband noise acting on their organs of hearing. It is shown that, under such conditions, the echolocating pulses of a dolphin become noticeably modified: in the absence of noise, standard broadband pulses are produced, while in the presence of noise, the pulses acquire an oscillatory character (become narrowband). Sounds and air pressure that occur inside the respiratory tract of a dolphin when the animal produces whistles and pulsed signals are studied. Data testifying in favor of the pneumatic origin of sounds generated by dolphins are obtained. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Mechanisms of radiation and reception of sound by dolphins represent the most complicated problems of the acoustics of cetaceans. In solving the problem of radiation, researchers still have to rely on indirect data obtained from the analysis of the far-field records of dolphin's sound signals and on the data obtained from morphological studies. Therefore, the nature of the sound source and its localization can only be subjects of speculation without any unique answer. The points where the acoustic information enters the head of a dolphin are also open to question. In addition to the evident assumption that the acoustic information is supplied to the internal ear of the dolphin through the auditory canals, other hypotheses have been put forward. In [1], it was assumed that, in cetaceans, sound is supplied to the cochlea through the lower jaw. This idea received further development in [2]. In another publication [3], it was assumed that an important role in the sound transmission to the cochlea is played by the frontal eminence. More reliable information on the mechanism of radiation and reception of sound could be obtained by an instrumental study of the expected area of the source and receiver localization. However, this is a rather difficult problem, which requires the development of special measuring instruments and a corresponding measuring technique. Nevertheless, such attempts have been made. The first successful experiments that made it possible to approach the source of sound of a dolphin consisted of recording the sound signals immediately on the dolphin's head [4, 5]. This method proved to be rather fruitful and revealed a number of characteristic features of sound radiation by dolphins [6–8, 9]. However, the possibilities offered by this method proved to be much wider owing to the development of a special set of instruments that can be fixed on a dolphin and allow a comprehensive study of its acoustic system. The set includes the following instruments: a three-channel broadband magnetic tape recorder, three miniature hydrophones, two independent noise generators, several noise radiators, a remote control system for controlling the operation of the tape recorder and the noise generators, and a tracer that provides a visual tracing of dolphin's motion in total darkness. Such a set of instruments makes it possible to study the echolocating behavior of a dolphin in situations complicated by the presence of artificially produced noise.

This paper describes some results obtained by studying the acoustic system of a dolphin (*Tursiops truncatus*) with the use of the aforementioned set of instruments.

EFFECT OF CORRELATED BROADBAND ACOUSTIC NOISE ON THE ECHOLOCATING BEHAVIOR OF A DOLPHIN

The experiment consisted of studying the effect of intense broadband correlated acoustic noise, which was presented to an adult dolphin in different local areas of its head (near the presumed acoustic inputs of the auditory system: near the auditory canals, at the lower jaw, and at the frontal eminence), on the echolocating activity of a dolphin while the latter solved the problem of detecting a fish presented to it in total darkness in an experimental tank. Correlated noise was produced by two or many radiators fed by a single generator. The basic idea of the experiment was that the echolocating abilities of dolphins should most fully manifest themselves under adverse conditions, when echolocation is hindered by external factors but serves as the only detection means. The experiment included two stages. At the first stage, two correlated noise radiators (spheres 30 mm in diameter, made of lead zirconate titanate piezoceramics) were fixed near the left and

1063-7710/04/5003-0343$26.00

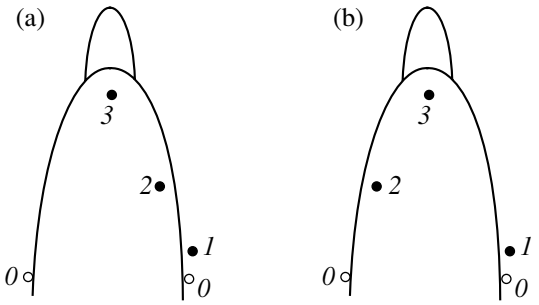


Fig. 1. Positions of (*0*) noise radiators and (*1–3*) hydrophones on the head of the dolphin: (a, b) different experimental versions.

right auditory canals (points *0* in Fig. 1a). Three hydrophones were fixed as follows: hydrophone *1* in the form of a piezoceramic sphere 15 mm in diameter was fixed near radiator *0* in the region of the right auditory canal; hydrophone *2* in the form of a piezoceramic cylinder 2 mm in diameter and 3 mm in height was fixed to the frontal eminence on its right-hand side, at approximately equal distances from the blowhole and the front edge of the frontal eminence; hydrophone *3*, identical to hydrophone *2*, was placed at the front edge of the frontal eminence, at the point where it joins the rostrum. The radiators were connected with the noise generator, and the hydrophones, with the inputs of the three-channel magnetic tape recorder. The noise generator and the tape recorder were fixed to the dorsal fin of the dolphin. The operation of the noise generator and the tape recorder was controlled through an underwater radio channel by radio signals with a duration of about 1 s.

At the second stage of the experiment, two correlated noise radiators were positioned at the lower jaw of the animal, at the left and at the right, and the third radiator was fixed to the frontal eminence exactly in the middle between the blowhole and the rostrum. The radiators were identical with those used at the first stage of the experiment. Hydrophones were positioned as shown in Fig. 1b and were also identical with those used at the first stage. In the course of the experiment, the dolphin under study swam in a tank with horizontal dimensions of 12.5×6 m and a depth of 1.2 m. The starting position of the dolphin was near one of the shorter walls of the tank. Near the opposite shorter wall, in one of the corners, a fish was presented to the dolphin. The fish was suspended from a thin thread and placed in water with a loud splash, which served as the starting signal. From this moment on, the dolphin began moving toward the fish while the fish, in its turn, was noiselessly moved over a distance of 2.5–3 m along the shorter wall of the tank. Since the experiment was performed in total darkness, the dolphin had to use echolocation. The distance between the starting position of the dolphin and the fish was 9–10 m. After the dolphin traveled approximately half the distance to the fish, broadband acoustic noise was turned on. The noise was produced by the aforementioned radiators fixed on the head of the dolphin. Its spectrum was characterized by a constant level within the frequency band from 5 to 30 kHz and a decrease of 6–7 dB/octave at higher frequencies. The noise pressure level in the total frequency band was 120 ± 6 dB.

The response of the dolphin to the appearance of noise was determined in two ways: first, visually, by the change in the trajectory of the luminous tracer and, second, by the change in the echolocation activity.

At the first stage of the experiment, when noise was presented in the region of the auditory canals, a pronounced motor response was observed in the form of an abrupt change in the tracer trajectory. The dolphin started at the onset of noise. Some experiments were carried out in the day time (in full light), and one could see that, when the noise was turned on, the dolphin strongly moved its head trying to free it from the noise source but soon calmed itself. At the second stage of the experiment, when the noise was presented near the lower jaw and the frontal eminence, no motor activity was observed. A similar situation was observed for the echolocating activity.

In the absence of noise, the dolphin detected underwater objects, including the fish presented to it, by using standard pulses of the type shown in Fig. 2a. The pulse shape may vary within certain limits [6, 9], but, on the average, all pulses are close in shape and have broadband spectra. The energy spectrum of these pulses is shown in Fig. 2c (curve *a*). In the literature, such pulses are considered as highly stable to varying experimental conditions and, in particular, to the appearance of noise radiation from distant sources [10]. Our experiments with the presentation of correlated noise in the region of auditory canals gave an unexpected result. At the moment when the noise was turned on, the dolphin, within a short time interval of 200–300 ms, performed a transformation of the pulse spectrum by making it much narrower. A standard (or close-to-standard) pulse was transformed to an oscillatory pulse (Fig. 2b), and its spectrum became narrowband (curve *b* in Fig. 2c). This transformation was systematically detected by hydrophone *3* (positioned near the rostrum) when noise was presented near the auditory canals. An even more unexpected result was that the pulses detected by hydrophone *2* (see Fig. 1a) remained standard, although they were synchronous with pulses detected by hydrophone *3*. However, the standard pulses detected by hydrophone *2* noticeably increased in amplitude (by a factor of 2.5–3) while the pulses detected by hydrophone *3* changed their spectrum but retained their amplitude. In some cases, hydrophone *2* fixed on one side of the frontal eminence of the dolphin also detected oscillatory pulses. In particular, an echolocation series consisting of alternating standard and oscillatory pulses spaced at several milliseconds was recorded. This series can be interpreted as a result

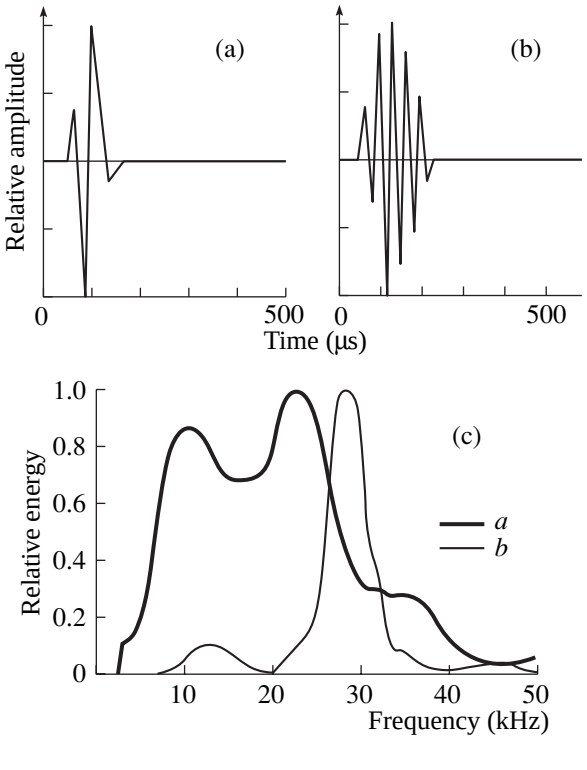


Fig. 2. (a) Standard and (b) oscillatory pulses and (c) their frequency spectra.

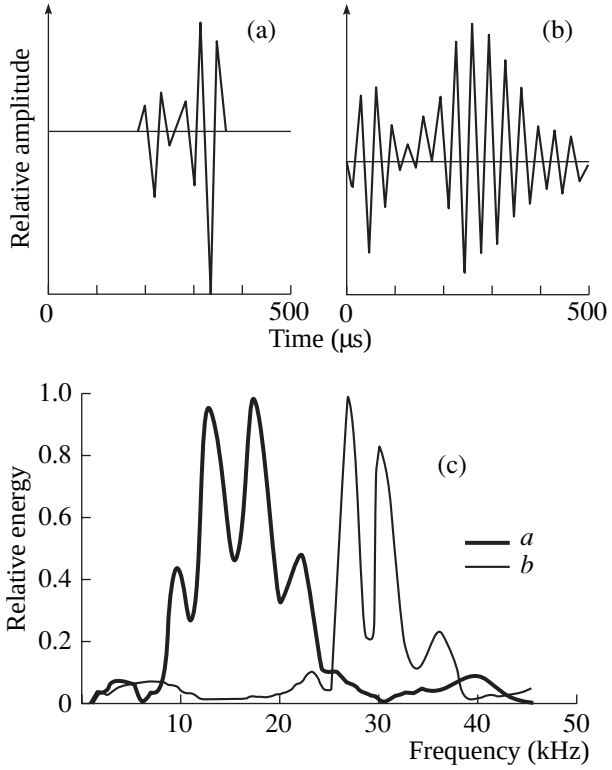


Fig. 3. (a) Standard and (b) oscillatory pulses with precursors and (c) their frequency spectra.

of the superposition of two simultaneously generated series, one of which consists of standard pulses and the other, of oscillatory pulses.

Note that, in the oscillograms of pulses recorded in the presence of noise, the latter is not observed, because the noise was produced near the auditory canals while the echolocation signals were recorded at the frontal eminence, where the ratio of the signal amplitude to the rms noise amplitude was no less than 40 dB.

Another important fact is as follows. The pulses produced by the dolphin in the presence of noise sometimes have "precursors," which were previously observed by Zaslavskii [11] and Dubrovsky [10]. By now, the presence of only one precursor before a pulse can be stated with confidence. The amplitude of such a precursor can vary over wide limits and may even reach the amplitude of the main pulse, so that the pulse becomes double. However, the duration of the precursor remains somewhat smaller than that of the main pulse. Oscillograms of pulses with precursors recorded by hydrophones *2* and *3* in the presence of noise are shown in Figs. 3a and 3b. Their energy spectra are shown in Fig. 3c. The characteristic feature of these spectra is their complex shape.

Hydrophone *1* positioned near the auditory canal never detected any oscillatory pulse radiation. Instead, at the same instants of time, it detected pulses close to standard ones. The oscillatory pulses may reach the hydrophone positioned near the auditory canal only in the form of pulses reflected from objects located before the dolphin.

At the second stage of the experiment, when the noise radiators were positioned at the lower jaw and at the frontal eminence of the dolphin, no considerable changes were observed in the echolocating activity. The pulses produced in the absence of noise and after the noise was turned on differed little from the standard (at the sites of hydrophones *2* and *3*). One can only state that the dolphin did hear the noise: a slight increase in the pulse amplitude was detected by all three hydrophones. At the same time, no noticeable changes were observed in the pulse shape and the pulse rate.

EFFECT OF UNCORRELATED BROADBAND ACOUSTIC NOISE ON THE ECHOLOCATING BEHAVIOR OF A DOLPHIN

In this experiment, another adult dolphin was used. Now, the fish detection task was even more complicated. To make the detection by echolocation most difficult, independent noise generators were used to produce an uncorrelated broadband acoustic noise field of high intensity near each of the auditory canals of the dolphin. The noise level was about 130 dB. Note that uncorrelated noise is produced by several noise radiators fed from different noise generators. In the case

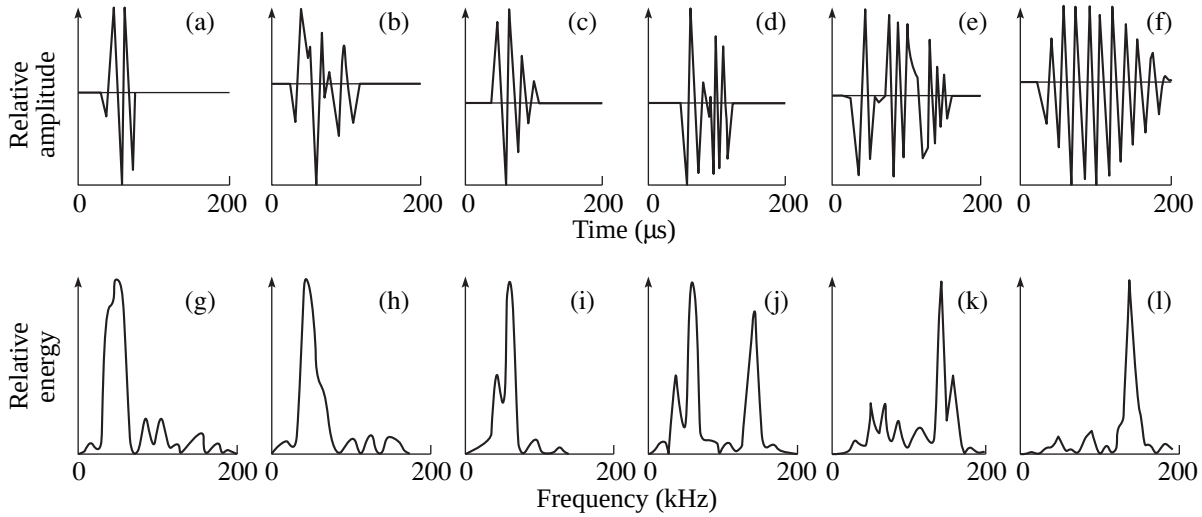


Fig. 4. Transition from standard to oscillatory pulses within one series: (a–f) pulses of the series and (g–l) their frequency spectra.

under consideration, the noise level was only 6–9 dB below the limiting intensity level at which the dolphin refused to locate the fish. The noise spectrum was uniform in the frequency band from 8 to 50 kHz, and it increased by 6–8 dB/octave up to a frequency of 180 kHz. The positions of hydrophones on the head of the dolphin are shown in Fig. 1b.

The main results can be formulated as follows.

The experiment with uncorrelated noise did not reveal such a pronounced dependence of time–frequency characteristics of dolphin's echolocation pulses on the presence or absence of noise as was observed in the experiment with correlated noise. In most cases, the dolphin preferred to use oscillatory pulses. It seems that, being frightened by intense uncorrelated noise at the first noise presentation, the dolphin continued using oscillatory pulses even during the time intervals when noise was turned off. Only from time to time, the dolphin introduced standard pulses between oscillatory pulses of a series. Another characteristic feature of dolphin's acoustic behavior was the fact that, in this case, oscillatory pulses usually contained 10–15 periods.

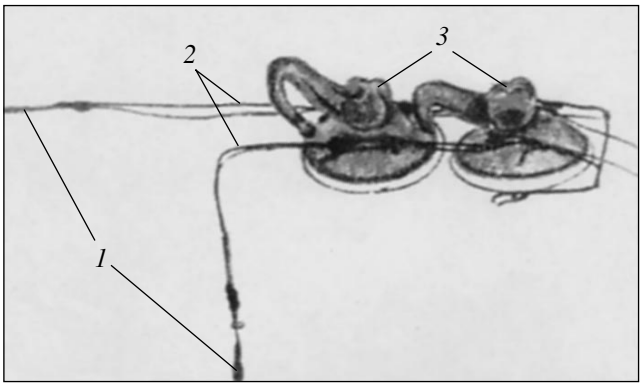


Fig. 5. Setup for introducing the hydrophones into the respiratory tract of the dolphin. Explanations are given in the main body of the paper.

Often, a smooth transition from standard to oscillatory pulse radiation could be observed. Figures 4a–4f show such a transition. The pulses were recorded by the hydrophone positioned near the rostrum. Figures 4g–4l show the energy spectra of these pulses. One can notice that, in the experiment with uncorrelated noise, the dolphin exhibited higher-frequency properties, as compared to the case of correlated noise.

Let us consider in more detail the pulse shown in Fig. 4d. It exhibits two pronounced features. The first consists in that the pulse begins with a fragment (precursor) similar to a standard pulse. An analysis of a great number of pulses showed that almost all of them begin with such fragments (a similar fragment can be seen in the pulse shown in Fig. 4e, although, here, it is less pronounced). The second feature of the pulse consists in that its higher-frequency part is separated into two sections (the pulse shown in Fig. 4f has no such feature). Presumably, such a separation is not incidental: we observed series of several tens of pulses that had low-frequency fragments (precursors) and were separated into two or even three sections. Pulses that were not separated into sections were observed much more rarely.

As in the experiment with correlated noise, the hydrophone positioned near the auditory canal never detected oscillatory pulse radiation. Instead, at the same instants of time, it detected pulses close to the standard ones. The oscillatory pulses reach the hydrophone near the auditory canal only as pulses reflected from objects located in front of the dolphin.

In the experiment with uncorrelated noise, the hydrophone positioned on one side of the frontal eminence almost always detected oscillatory pulses simul-

taneously with the hydrophone positioned near the rostrum. By contrast, in the experiment with correlated noise, such observations were very rare: the side hydrophone almost always detected standard pulses.

The results presented above suggest the following conclusions:

—dolphins are capable of changing on purpose the time–frequency characteristics of their echolocation pulses under the effect of intense broadband noise presented in the area of their auditory canals;

—it is possible that dolphins possess different sources for producing standard and oscillatory pulse radiation;

—acoustic information is supplied to the internal ear of dolphin's auditory system through the auditory canals; and

—the system of oscillatory pulse radiation has a low *Q* factor and operates in a forced operating mode.

SOUNDS INSIDE THE RESPIRATORY TRACT OF A DOLPHIN

The set of instruments described above allows one to penetrate into the respiratory tract of a dolphin and to detect the sounds that accompany the radiation of echolocating pulses. This kind of information should be useful in discussing the hypothesis that the signals produced by a dolphin are of pneumatic origin [12]. For recording the sounds inside the respiratory tract of a dolphin, the miniature broadband hydrophones described above (hydrophones *2* and *3* from Fig. 1) were used. They were mounted on special holders shown in Fig. 5. The hydrophones (objects *1* in Fig. 5) were fixed in the required position by (*2*) rigid holders made of a stainless steel wire 1 mm in diameter. The holders, in their turn, were attached to (*3*) two suction cups used to fix the whole system on the head of the dolphin under test. The setup shown in Fig. 5 is intended for introducing one hydrophone into the blowhole of the dolphin (the holder of this hydrophone is bent at right angle). The other hydrophone is positioned above the head of the animal, before the blowhole. In the experiment described here, the hydrophone was introduced in the blowhole to a depth of 6 cm. The process of introducing the hydrophone seemed not to annoy the animal to any considerable extent: after two or three attempts to push the hydrophone out of the blowhole by an air flow, the dolphin resigned itself to the situation and became willing to locate and take the fish. Figure 6 illustrates the process of introducing the hydrophone into the blowhole. In the same figure, one can see the noise radiators and other hydrophones fixed on the head of the animal.

The miniature hydrophones were alternatively introduced into the right and left nostrils. During the periods of active echolocation, which was detected by the hydrophones positioned on the head of the dolphin, the sound pulses inside the nostrils were either absent or much weaker than outside. This suggests that the larynx is not the source of echolocation signals, as was presumed by Gurevich [13, 14]. In addition, in the right nostril, the inner hydrophone detected additional sounds that could not be heard outside but accompanied the echolocation. These sounds were similar to whistles (or mewing) in character and noticeably differed from communication whistles in their structure. As a rule,

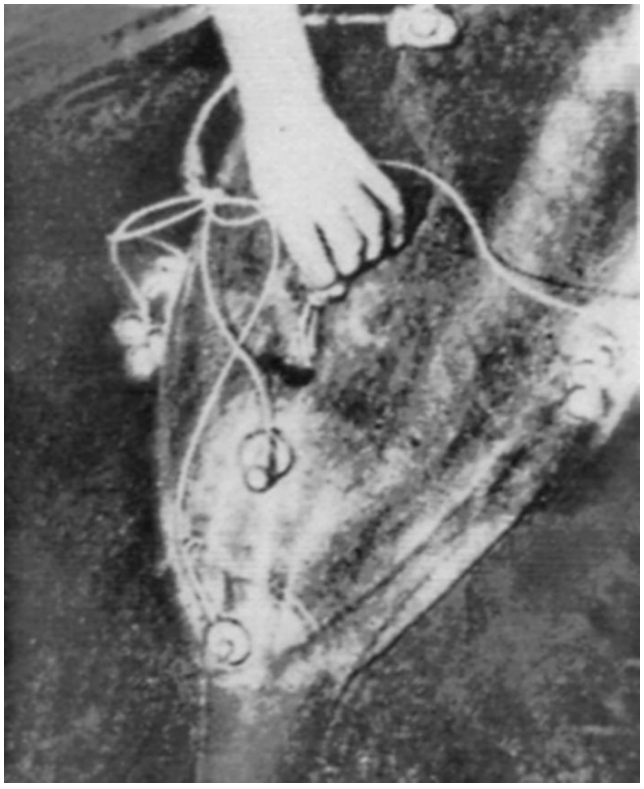


Fig. 6. Introduction of the hydrophone into the respiratory tract of the dolphin.

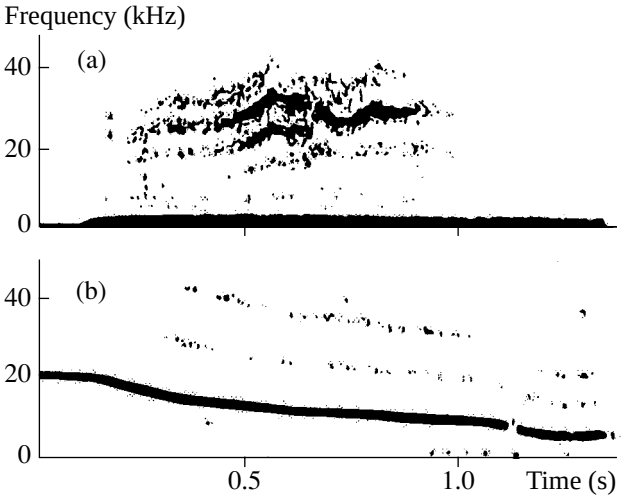


Fig. 7. Sonograms of an (a) internal sound and (b) communication whistle.

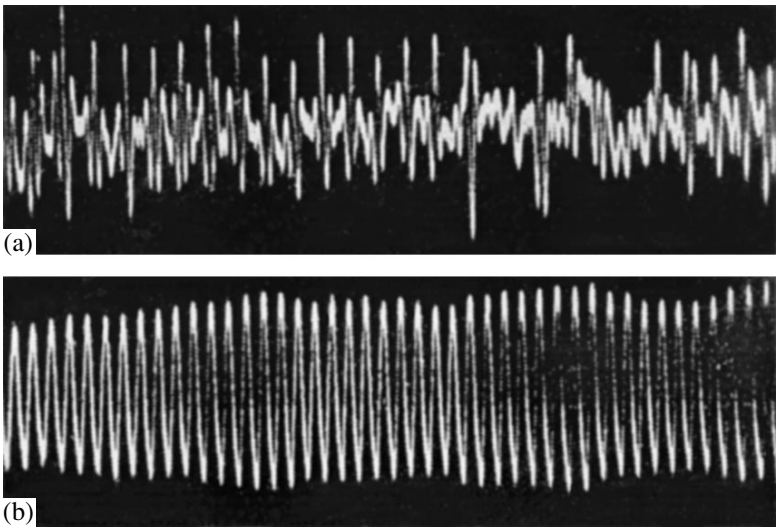


Fig. 8. Oscillograms of (a) internal sound and (b) communication whistle.

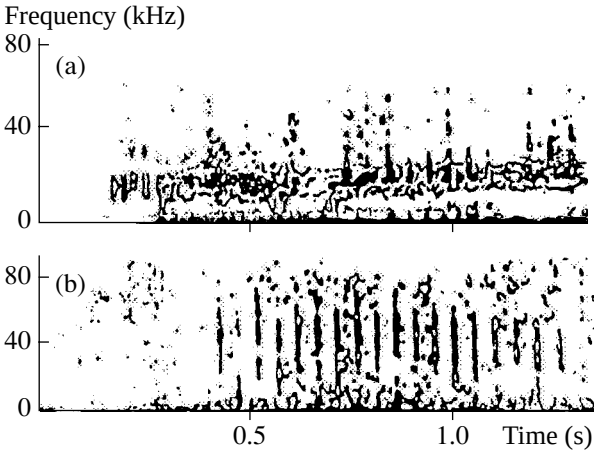


Fig. 9. Sonograms of (a) internal sound on the background of a pulse series and (b) the same pulse series recorded on the head of the dolphin.

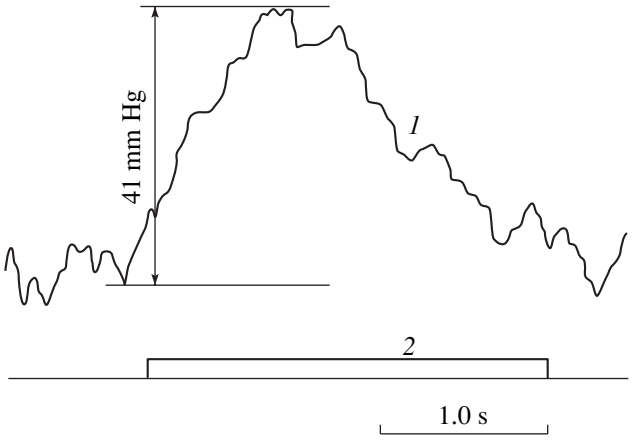


Fig. 10. Air pressure inside the nostril of an echolocating dolphin. (For explanations, see the main body of the paper.)

they had the form of narrowband sounds, sometimes with a pronounced amplitude modulation characterized by side frequencies that were not multiples of the fundamental frequency (Fig. 7a). Communication whistles usually contain multiple harmonics (Fig. 7b). Figures 8a and 8b show the oscillograms of an internal sound and a communication whistle. One can see that the latter is of a mainly harmonic character while the internal sound has the form of a sequence of one- and two-period pulses, which can be interpreted as a signal with amplitude overshooting. Figures 9a and 9b present the sonograms of internal sound on the background of a weakly noticeable series of pulses (inside the blowhole) and of the same series of echolocation pulses recorded on the head of the dolphin, near the rostrum. It should be stressed that all internal sounds described above were detected in the right nostril. In the left nostril, no sounds were observed. Unfortunately, it was impossible to repeat many times the attempts to detect any sound in the left nostril to reliably determine whether internal sounds do or do not occur there in the course of echolocation. The limited amount of data available to us allows only a preliminary conclusion that the left nostril area is below the muscle plug and has nothing to do with echolocation while the right nostril is evidently involved in the process.

Practically all internal sounds recorded in our experiment coincide in time with echolocation. Only in two cases out of seventeen, internal sounds were not accompanied by echolocation (possibly, the latter was very weak and could not be detected). In other two cases out of the same seventeen, internal sounds were not detected in the course of echolocation (possibly, they were very weak). The duration of internal sounds does

not coincide with the duration of echolocation series: as a rule, the former are shorter. Possibly, internal sounds are of an aerodynamic origin and accompany the air flow through the structures responsible for the generation of echolocation sounds. Another explanation is also possible: internal sounds may be caused by a considerable difference between pressures on the two sides of the muscle plug and be generated as a result of the air leakage through the muscle plug. This explanation is confirmed by the pressure measurements in the respiratory tract of dolphins (Fig. 10) [15]. One can see that the air pressure in the nostril between the inner muscle plug and the outer valve (curve *1* in Fig. 10) increases in the course of echolocation (line *2*) by 30–50 mm Hg. Subsequent studies should elucidate the nature of internal sounds. As of now, the presence of these sounds testifies in favor of the pneumatic mechanism of sound generation by dolphins.

REFERENCES

1. K. S. Norris, in *Marine Bioacoustics*, Ed. by W. N. Tavolga (Pergamon Press, Oxford, 1964), pp. 317–336.
2. L. K. Rimskaya-Korsakova and N. A. Dubrovsky, Sens. Sist. **12** (5), 496 (1998).
3. J. G. McCormick, E. G. Wever, J. Palm, and S. H. Ridgway, J. Acoust. Soc. Am. **48**, 1418 (1970).
4. K. J. Diercks, R. T. Trochta, C. F. Greenlow, and W. E. Evans, J. Acoust. Soc. Am. **49**, 1729 (1971).
5. E. V. Romanenko, in *Proceedings of V All-Union Workshop on Biology of Marine Mammals* (Makhachkala, 1972), p. 73.
6. E. V. Romanenko, *Physical Fundamentals of Bioacoustics* (Nauka, Moscow, 1974).
7. E. V. Romanenko, in *Marine Mammals: Results and Methods of Hydrobionic Studies* (Nauka, Moscow, 1977), pp. 157–168.
8. E. V. Romanenko, in *Proceedings of IX All-Union Acoustical Conference* (Moscow, 1977), Sect. Ts, p. 49.
9. C. Kamminga, PhD Thesis (Delft Univ. of Technology, Delft, 1994).
10. N. A. Dubrovsky, *Dolphins' Echolocation (a Review)* (TsNII Rumb, Leningrad, 1975).
11. G. L. Zaslavskiĭ, Candidate's Dissertation in Biology (Karadag, 1974).
12. É. Sh. Aĭrapet'yants and A. M. Konstantinov, *Echolocation in Nature* (Nauka, Leningrad, 1974).
13. V. S. Gurevich, in *Application of Mathematical Methods and Computer Technology in Medicine and Biology* (Nauka, Leningrad, 1972), pp. 123–125.
14. V. S. Gurevich, in *Proceedings of V All-Union Workshop on Biology of Marine Mammals* (Makhachkala, 1972), p. 34.
15. V. G. Dargol'ts, E. V. Romanenko, E. A. Yumatov, and V. G. Yanov, Fiziol. Zh. **67** (11), 1744 (1981).

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 350–358. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 419–428.

BIOLOGICAL ACOUSTICS

Binaural Unmasking of the Periodic Component in the Envelope of an Amplitude-Modulated Signal

V. N. Telepnev

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

e-mail: telepnev@akin.ru

Received October 12, 2003

Abstract—A binaural unmasking of a tone component that is present in an amplitude–time noise envelope of a high-frequency signal is studied. The signal has the form of a sinusoidal carrier of frequency 2000–5000 Hz amplitude modulated by a low-frequency signal. The modulating function is a mixture of a 300-Hz tone (interaurally inphase or antiphase) and a dichotic masking noise within 0–400 Hz, this mixture being subjected to a half-wave linear rectification. The listener has to detect the rhythmic component in the modulating noise function. It is shown that, under the aforementioned conditions, the binaural difference in masking levels grows up to 25 dB with increasing carrier frequency but drastically decreases in the case of a masking of the low-frequency part of the basilar membrane in the vicinity of 300 Hz. The lateralization based on the interaural phase of a 100% amplitude modulation by a 300-Hz tone at a carrier frequency within 2000 to 5000 Hz also drastically decreases (in our experiments) when the low-frequency part of the basilar membrane is masked. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is well known that the masking threshold (level) is maximal when the interaural phase differences of a desired tone signal and the masker are identical. If the interaural phase difference of a tone signal is varied, the masking threshold decreases. This testifies to the presence of a binaural unmasking, which is characterized by the difference in the masking levels of interaurally inphase and antiphase tone signals in interaurally coherent noise, or by the binaural masking level difference (BMLD). For low-frequency tone signals (below 500 Hz) masked by broadband noise, the BMLD reaches 15 dB. For high-frequency tones (above 2000 Hz), the BMLD does not exceed 3 dB [1]. In the case of complex signals, the BMLD vanishes when the signals contain no low-frequency components. This occurs because the reproduction of the temporal profile of the signal in the firing rate of auditory neurons is possible only for low-frequency signals. For high-frequency signals, only the low-frequency envelope of sound is retained in the auditory nerve.

However, when masking by narrowband noise takes place, a binaural unmasking is also possible for high-frequency tones [2–4]. This is explained by the fact that narrowband Gaussian noise has a deep low-frequency envelope, which is modified when the tone is added to the noise masker and which depends on the temporal shifts of the tone. As a result, the addition of an interaurally antiphase tone leads to an interaural decorrelation of low-frequency envelopes and facilitates the detection of the interaurally antiphase tone. At first it was believed that, under these conditions, the BMLD of high-frequency tones also reaches 15 dB. However, subsequent experiments have shown that the BMLD grows with an increase in the peak-factor of the envelope (the ratio of the rms value to the mean value) [5–7] but does not exceed 8–10 dB under conventional experimental conditions.

In our previous publications [8–12], we reported on the binaural unmasking of a periodic component of the amplitude envelope of a high-frequency signal, which had the form of a sine carrier in the frequency band from 2000 to 5000 Hz with an amplitude modulation (AM) by a low-frequency mixture of a tone signal and a masker. It was shown that, under such conditions, the BMLD may reach 30 dB. The present paper reviews the results reported in the cited publications and describes the results of the latest experimental studies.

EXPERIMENTAL CONDITIONS AND SIGNALS

Experiments were planned to be performed with high-frequency signals that retain the information on their phase after their reception. The experiments with amplitude-modulated signals were expected to be of interest because of the possibility to control the modulation depth. The whole series of experiments was planned to be performed with the same group of listeners, no matter how small this group is. The listener had to detect the interaurally inphase or antiphase rhythmic (tone) component in interaurally coherent narrowband noise envelope of an amplitude-modulated signal. The modulating function was a mixture of a tone (interaurally inphase or antiphase) of frequency 100–300 Hz and a masking noise (interaurally inphase) in the fre-

1063-7710/04/5003-0350$26.00 © 2004 MAIK "Nauka/Interperiodica"

quency band from 0 to 400 Hz. These experimental conditions will be referred to as "linear AM conditions."

In 1998–2000, we have found new experimental conditions of binaural unmasking of high-frequency signals. In these experiments, the amplitude envelope of the high-frequency tone carrier was formed so as to reproduce the form of the receptor potential and the poststimulus histogram of low-frequency neurons if the stimulation were directly performed by the low-frequency oscillation used in the experiment as the envelope. For this purpose, the modulating function was chosen to be in the form of a low-frequency mixture of the signal (a 300-Hz tone) and the noise masker, this mixture being subjected to a half-wave detection (rectification). The role of the modulator was played by a multiplier. One would expect that, under these experimental conditions, which will be referred to as "the half-wave AM conditions," the receptor potential and the poststimulus histogram of high-frequency neurons (with characteristic frequencies near the carrier frequency of 3000 Hz) in the given experiments should be qualitatively similar to those of low-frequency neurons with characteristic frequencies near 300 Hz in the experiments demonstrating the binaural unmasking of a 300-Hz tone in low-frequency noise.

Figure 1 shows examples of realizations of experimental signals. The periodic component (300 Hz) of the envelope is sufficiently pronounced: one can clearly distinguish five periods. Figures 1a and 1b represent the amplitude-modulated oscillation and its spectrum, and Figs. 1c and 1d show the corresponding characteristics obtained with a preliminary half-wave rectification of the modulating function.

For control and comparison, we also estimated the BMLD under conventional experimental conditions with masking by broadband and narrowband noise.

METHOD

Signals were generated by a computer program via an SB-16 sound card at a sampling rate of 44100 Hz. The same program processed the responses of the listener and controlled the experiment.

The test sequences consisted of two or four signals. In the latter case, the first and fourth signals contained the masker alone. The listener should determine which of the two signals (second or third) contained a tone added to the masker. The pause between the signals of the test sequence varied from 0.7 to 1.0 s for different experiments. The masker duration was 0.5 s, and the tone signal was 20 ms shorter than the masker and was centered with respect to it. The time of linear build-up and fall-off was 5 ms for both masker and signal.

The masking thresholds were estimated using the adaptive procedure of two-alternate forced choice by the following scheme: after each error the signal was increased by one step, and after two consecutive correct responses, the signal was decreased by one step. These signal variations were performed in the modulating function under the linear AM and half-wave AM conditions. The masker realization was not varied within each single BMLD estimation process. At the beginning of the experiment, the signal amplitude was chosen so as to provide the signal detection with confidence, and the step was taken to be 30% of the current amplitude. Every time the current rms deviation of the turning points became smaller than one step, the step was reduced by half. However, it never became smaller than 7.5%.

The first two turning points were ignored. In the first half of the study, the detection threshold was determined by averaging over 16 turning points. In one experiment, the standard deviation of the threshold estimate, which was determined from the turning points, was usually found to be smaller than 1 dB. However, the difference between estimates provided by replicate experiments could exceed 10 dB, which is common to psychoacoustics. In addition, the comparison of the threshold estimates by 8 and 16 turning points showed that, in the latter case, the thresholds were somewhat higher (presumably because of the fatigue of the listener) while their scatter was no smaller than in the first case. The threshold estimates in the second half of the study were obtained from 8 turning points.

In the course of a single experiment, first, the masking level was estimated by 8 turning points for an interaurally inphase (antiphase) tone masked by a single (recorded and repeated) realization of the noise masker. Then, the listener, without leaving the acoustic cabin, estimated the masking level by 8 turning points for an interaurally antiphase (inphase) tone masked by the same realization of the noise masker. Thus, for one BMLD estimate, one masker realization was used. The initial phase of the tone signal in each test was chosen randomly according to the uniform distribution over the interval from 0 to 2π. In this manner, the phase–time relations between signal and masker were randomized.

As a masker, we used interaurally inphase noise formed in one of the two following ways: first, by filtering (using the computer program) the sequence obtained from a random number generator by a low-frequency or bandpass filter of second order, and, second, by a sum of 50–100 sinusoids with a Rayleigh amplitude distribution and a uniform phase distribution, which is a model of normal noise.

The adaptive procedure of two-alternate forced choice was also used for estimating the lateralization threshold based on the interaural phase differences of the modulating tone function at a 100% amplitude modulation. Two signals of the test sequence received interaural phase shifts of the sine amplitude envelope so that these phase shifts were equal in magnitude and opposite in sign. One of these shifts corresponded to a shift of the auditory pattern to the right and the other, to the left. The listener had to determine which of the sig-

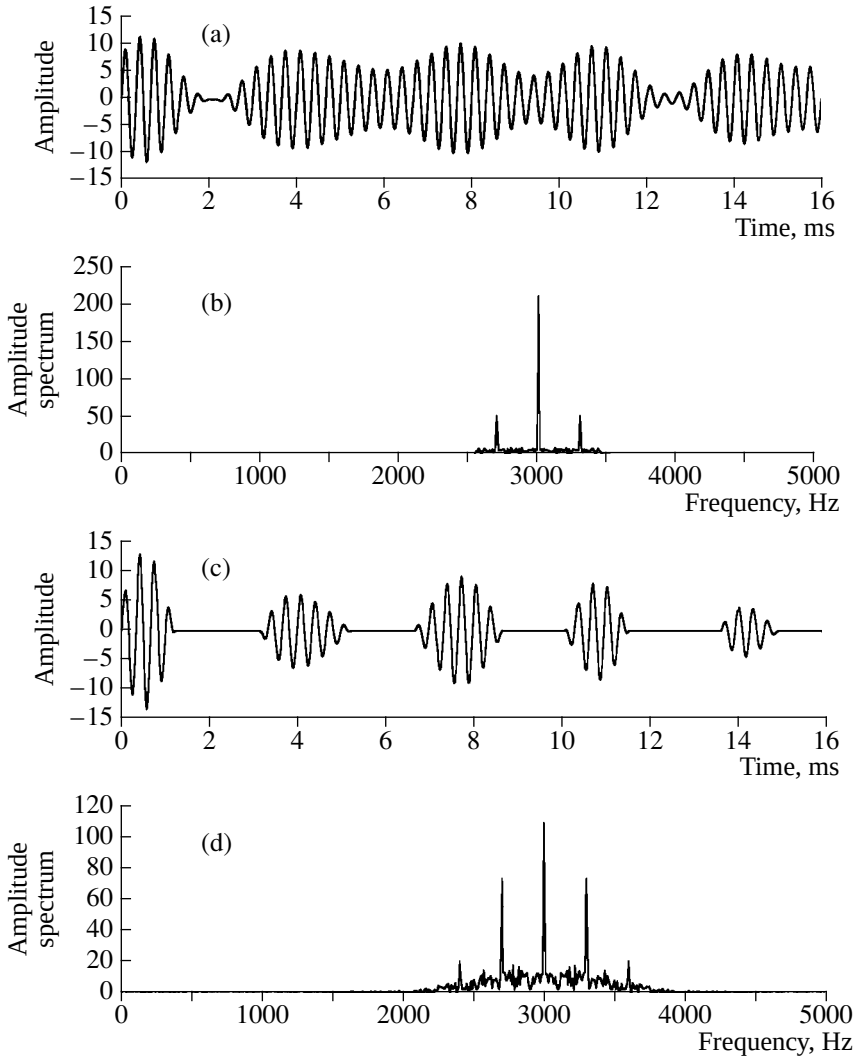


Fig. 1. (a, c) Time realizations of signals and (b, d) their spectra: (a, b) linear AM conditions of the experiment and (c, d) half-wave AM conditions.

nals (first or second) was shifted to the right. After each error, the interaural delay (phase) of the envelopes was increased by 30%, and after two consecutive correct responses, it was decreased by 30%.

The adaptive procedure does not always lead to success if the lateralization threshold is estimated under the effect of noise. For such experimental conditions, the statistical method was used. In this case, the phase step of the 100% amplitude modulation corresponded to a delay of 80 μs. The set of phase shifts consisted of ten values corresponding to delays from 80 to 800 μs. One of the signals of the experimental sequence received a phase shift corresponding to lateralization to the right, and the other signal received a phase shift of the same magnitude that corresponded to lateralization to the left. The phase shifts were chosen by the computer program at random but so as to use all ten possible values from 80 to 800 μs in ten tests. In each test, the listener indicated which of the signals (first or second) was heard on the right. A series of ten tests was repeated ten times, so that one estimate of the lateralization threshold was obtained from 100 tests.

RESULTS

First, we estimated the BMLD of low-frequency and high-frequency tone signals by reproducing the experiments known from the literature [1, 2]. This was accomplished with participation of three listeners with normal hearing, 40 to 60 years old. In these experiments [8–10], the mean values of the BMLD of a 3000-Hz tone signal masked with narrowband noise did not exceed 5–10 dB for all three listeners.

To obtain additional information on the role of interaural decorrelation of low-frequency amplitude envelopes, we estimated the BMLD for the case of the detection of a 300-Hz sine component in low-frequency noise in a frequency band of 0–400 Hz, which modu-

lated a high-frequency carrier: the listener detected the rhythmic component in the modulating noise function. Under the linear AM conditions, at a carrier frequency of 3000 Hz, the BMLD proved to be approximately the same as the BMLD of a tone of the same frequency in a narrow-band masker.

Under the half-wave AM conditions, the BMLD drastically increased. All listeners noted that, in the case of interaurally antiphase signals, the auditory pattern contained pronounced right and left parts in addition to the central part. It is this subjective feature that reduces the detection threshold of an antiphase tone in the half-wave AM envelope. This property, although to a very small extent, is also present under the linear AM conditions, as well as under conventional conditions of detecting an interaurally antiphase tone in narrowband high-frequency noise.

The data obtained demonstrate an increase in the binaural unmasking with increasing carrier frequency and, presumably, testify to the relation between this phenomenon and the simultaneous increase in the critical bandwidth. According to experimental estimates [8–10] for the detection of a 300-Hz tone signal, at the carrier frequencies of 3000, 4000, and 5000 Hz, the BMLD was found to be equal to 13, 20, and 24 dB, respectively.

Figure 2 presents the dependences of the BMLD on the frequency of the detected tone signal in the modulating function and on the carrier frequency under the half-wave AM conditions for the listener called TVN. The masker in the modulating function was an imitation of normal noise of 100 tones spaced at 1 Hz.

One can assume that the BMLD reaches its maximal values when the envelope of the AM oscillation passes through the filter of the basilar membrane without distortions. In this case, the receptor potential and the post-stimulus histogram of high-frequency neurons with characteristic frequencies near the carrier frequency in the given experiments, presumably, coincide with the receptor potential and the post-stimulus histogram of low-frequency neurons that have characteristic frequencies near the frequency of the tone to be detected (70, 120, 200, or 300 Hz) in conventional experiments on detecting a low-frequency tone in narrowband noise. When this condition is satisfied, the receptor potential and the post-stimulus histogram have the maximum possible peak-factor. To retain the shape of the envelope after the AM signal passes through the filter of the basilar membrane, it is necessary that the critical bandwidth at the carrier frequency exceeds the AM signal spectrum width, which, under the half-wave AM conditions, is three to four times greater than under the linear AM conditions (Fig. 1). For example, at a modulation frequency of 120 Hz, the critical bandwidth should be 800–1000 Hz. Hence, one can expect an increase in the BMLD when the carrier frequency increases up to 4000–5000 Hz. Analogously, under the half-wave AM conditions, the BMLD decreased to 12 dB and did not

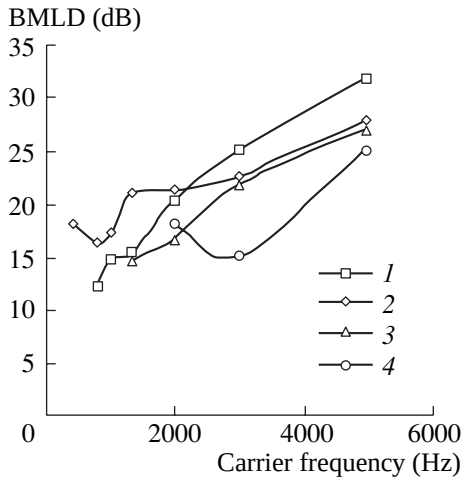


Fig. 2. Dependence of the BMLD on the carrier frequency under the half-wave AM conditions (the TVN listener). Each point is averaged over five experiments. The SPL is 80–85 dB. The tone frequency in the modulating function serves as the parameter of the curves: (*1*) 70, (*2*) 120, (*3*) 200, and (*4*) 300 Hz.

depend on the carrier frequency if, after a linear-response detector and before the modulation, the modulating function passed through a low-frequency filter of the second order with a cutoff frequency of 400 Hz and an attenuation of 12 dB per octave away from the cutoff frequency.

Another hypothesis is based on the assumption that, because of the nonlinearity of the basilar membrane, low-frequency mechanical oscillations occur in the internal ear, and these oscillations correspond to the amplitude–time envelope of the high-frequency signal. These oscillations excite the low-frequency part of the basilar membrane, exactly where the detection of the low-frequency sine component of the high-frequency signal envelope takes place. As a consequence, under the half-wave AM conditions, the BMLD typical of low-frequency signals is observed. To verify this assumption, the experiments described above were repeated with the same listeners, but now the low-frequency part of the basilar membrane was masked by independent low-frequency noise occupying the same spectral band as the modulating masker. For one of the listeners, at sound pressure levels (SPL) within 70–80 dB, the results were as follows: without masking of the low-frequency band, the BMLD exceeded 15 dB, and with masking of the low-frequency band, the BMLD was less than 4 dB. For the second listener, the BMLD was also found to decrease in the presence of masking of the low-frequency band, but, in this case, the decrease was smaller. Figure 3 shows the experimental estimates of the BMLD under the linear AM and half-wave AM conditions.

Simultaneously, with the same listeners, lateralization based on interaural amplitude differences and interaural phase differences of the amplitude envelope

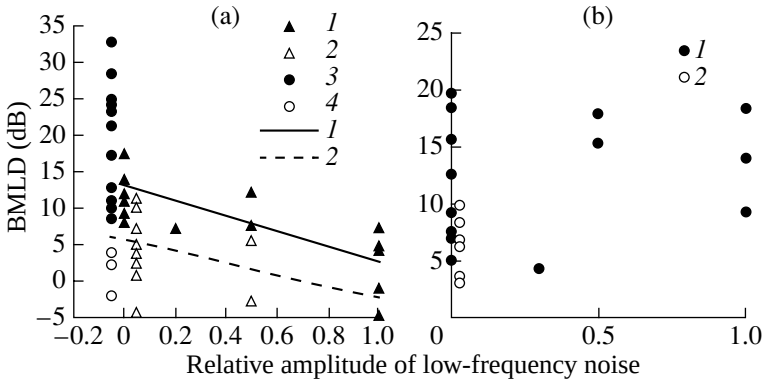


Fig. 3. BMLD in the amplitude envelope under the linear AM conditions and the half-wave AM conditions for two listeners. The SPL is 70–80 dB. The relative amplitude of low-frequency noise is equal to the ratio of its rms amplitude to the rms amplitude of the modulating masker. The full dots refer to the half-wave AM conditions, and the empty dots, to the linear AM conditions. (a) Data obtained with the TVN listener: (*1*) the half-wave AM conditions with a carrier frequency of 3000 Hz (17 points and the solid trend line); (*2*) the linear AM conditions with a carrier frequency of 3000 Hz (14 points and the dashed trend line); (*3*) the half-wave AM conditions with a carrier frequency of 5000 Hz (12 points); (*4*) the linear AM conditions with a carrier frequency of 5000 Hz (3 points: in this particular case, the points lying near the ordinate axis are displaced to separate them visually while in fact they lie on the ordinate axis, i.e., all of them correspond to the absence of low-frequency noise). (b) Data obtained with the TMV listener: (*1*) the half-wave AM conditions and a carrier frequency of 3000 Hz (14 points); (*2*) the linear AM conditions and a carrier frequency of 3000 Hz (6 points).

(a 100% sine amplitude modulation with a frequency of 300 Hz) was measured for 3000-Hz tone signals with masking of the low-frequency band of the basilar membrane by low-frequency noise. The measurements were performed using both the adaptive procedure of two-alternate forced choice and the statistical method. The relative amplitude of low-frequency noise was estimated as the ratio of its rms amplitude to the rms amplitude of the carrier.

Briefly, the results of these measurements can be summarized as follows:

—The measurements of the threshold values of the interaural delay (recalculated by the interaural phase difference) of the amplitude envelope were performed at the SPL within 70–80 dB.

—As can be seen from Figs. 4, 5, the lateralization thresholds based on the interaural phase differences of the amplitude envelope prove to depend (more or less) on the masking of the low-frequency band for both listeners.

Without masking of the low-frequency band of the basilar membrane, these thresholds are equal to 200–700 μs for both listeners.

When the low-frequency band is masked by noise with a spectral density from 40 to 60 dB, the lateralization as a rule is impossible, or the threshold interaural delay exceeds 1000 μs (the maximal measurable interaural delay is equal to the half-period of the AM frequency and is 1666 μs).

—The results are considerably scattered, in particular, because of the changes in the degree of concentration of the listener.

—The lateralization thresholds based on interaural amplitude differences were found to be independent of the masking of the low-frequency band of the basilar membrane for both listeners.

To make sure that the results obtained are not reduced to a simple masking of the frequency band of 3000–5000 Hz by low-frequency noise of frequencies 250–350 Hz, it is sufficient to consider the dependences of masking under the half-wave AM conditions at a carrier frequency of 5000 Hz on the amplitude and mean frequency of the second narrowband masker (Fig. 6). Remember that the detection of the 300-Hz sine component in the noise modulating function was performed in the presence of two independent maskers: one was present in the modulating function and the other was masking the given frequency band on the basilar membrane. Each of these maskers was an imitation of normal noise of 100 tones spaced at 1 Hz. The relative amplitude of the second masker took on only two values in these reference experiments: 0 and 0.3. From Fig. 6, one can see that the masking of the interaurally inphase periodic component in the amplitude modulation by the second masker is absent or is negligibly small compared to the masking by the modulating noise under the half-wave AM conditions, because the detection thresholds are practically independent of the amplitude and frequency of the second masker.

By contrast, the detection thresholds of interaurally antiphase periodic component in the modulating function strongly depend on the second masker, the masking being maximal when the mean frequency of the second masker is 300 Hz and almost absent (within the scatter of data) when the mean frequency of the second masker is 5000 Hz. One can assume that this is explained by the drop of the amplitude–frequency characteristic of the headphones at high frequencies: as a result, the masking of the frequency region near

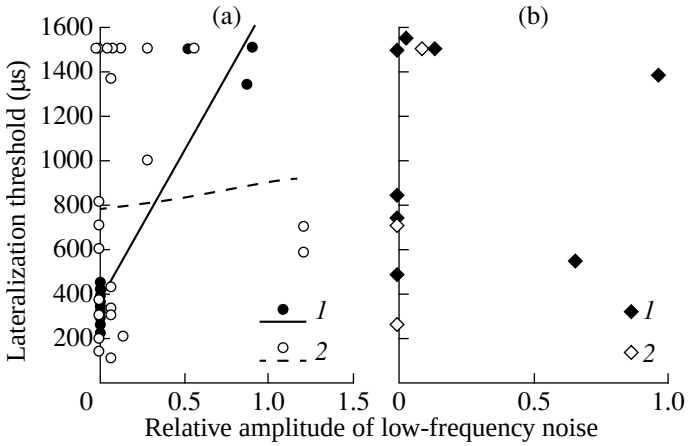


Fig. 4. Effect of low-frequency noise on the lateralization based on interaural phase differences of the amplitude–time envelope of a 100% AM for two listeners. The SPL of the carrier is 70–80 dB. The relative amplitude of noise is equal to the ratio of its rms amplitude to the rms amplitude of the carrier. (a) Data obtained with the TVN listener using the adaptive procedure of forced choice: (*1*) year 1998 (16 points and the solid trend line); (*2*) year 1999 (24 points and the dashed trend line). (b) Data obtained with the TMV listener (*1*) using the adaptive procedure of forced choice (8 points) and (*2*) using the statistical method (3 points).

5000 Hz by noise in a frequency band near 300 Hz is more efficient than by noise in the vicinity of 5000 Hz. To exclude this possibility, for the same listener we measured the absolute hearing thresholds at the frequencies of interest and the tone-by-tone masking thresholds at a frequency of 300 Hz by the conventional measuring technique (Fig. 7). For comparison, Fig. 7 also shows the mean hearing thresholds by Zwicker and Feldtkeller [14]. One can see that, at frequencies of 3000 and 5000 Hz, the listener exhibits a noticeable hearing loss, while the masking by a 300-Hz tone is almost absent at these frequencies.

Taking into account that, under the half-wave AM conditions, the nonlinear transformation consisting of the half-wave rectification leads to a broadening of the spectrum of the modulated signal, we studied the dependence of the BMLD on the bandwidth of both the masker spectrum and the spectrum of the signal to be detected under conventional experimental conditions [11, 12]. The results of these studies show that, in experiments with a simple increase in bandwidth, for example, in the case of the detection of a harmonic series including the frequencies of 2100, 2400, 2700, 3000, 3300, 3600, and 3900 Hz in noise whose spectrum lies within 1800–4100 Hz, the BMLD does not exceed 4 dB, as well as in the case of detecting a 3000-Hz tone in broadband noise. A similar result was obtained earlier [13].

At the same time, in the case of detecting a harmonic series of 2100, 2700, 3300, and 3900 Hz on the background of a masker, which has the form of a sum of a noise signal with a spectrum within 1800–2400 Hz and three noise signals obtained by heterodyning the first masker upwards by 600, 1200, and 1800 Hz, the BMLD increases to approximately 8 dB.

In addition, under conventional experimental conditions, we performed an experiment with a masker that was obtained by multiple heterodyning as described above but with a spectral overlapping of heterodyned parts of the masker. Such an overlapping occurs when the noise spectrum broadens under the half-wave AM conditions (Fig. 1). The configuration shown in Fig. 8 was formed by heterodyning the noise occupying the frequency band of 200–400 Hz with the help of a poly-

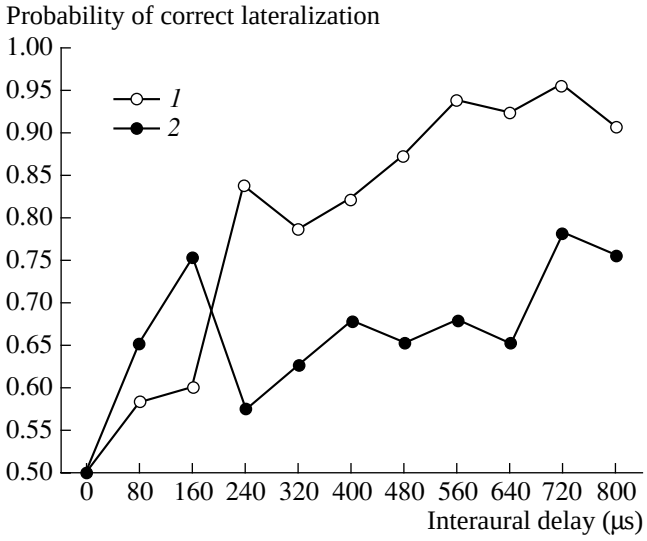


Fig. 5. Effect of low-frequency noise on the lateralization based on interaural phase differences in 100% amplitude modulation (the TVN listener with the statistical method). The SPL of the carrier is 80 dB. The relative amplitude of noise is equal to the ratio of its rms amplitude to the rms amplitude of the carrier. Tests (*1*) without noise (each point is a result of 60 tests) and (*2*) with a relative noise amplitude of 0.1 (each point is a result of 40 tests).

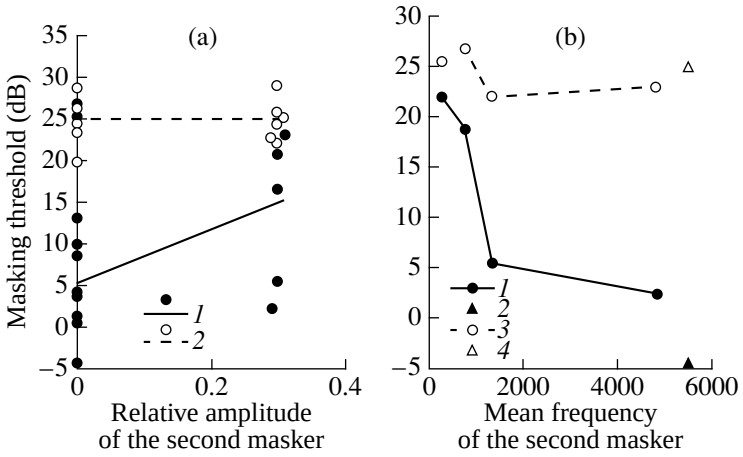


Fig. 6. Detection thresholds for the 300-Hz sine component in the modulating noise function under the half-wave AM conditions at a carrier frequency of 5000 Hz versus the amplitude and the mean frequency of the second narrowband (100 Hz) masker. (Data obtained with the TVN listener.) The SPL is 80 dB. The relative amplitude of the second masker is equal to the ratio of its rms amplitude to the rms amplitude of the modulating masker. (a) Dependence of the detection threshold on the amplitude of the second masker with a mean frequency of 300 Hz for (*1*) an antiphase signal (the solid trend line) and (*2*) an inphase signal (the dashed trend line). (b) Dependence of the detection threshold on the mean frequency of the second masker with an amplitude of 0.3 for (*1*) an antiphase signal, (*2*) an antiphase signal without masking, (*3*) and inphase signal, and (*4*) an inphase signal without masking.

harmonic carrier. The envelope of the time realization of the masker has a high peak-factor, but in this case, as before, the BMLD of the polyharmonic signal increases to approximately 8 dB.

DISCUSSION

The main results of the study are in qualitative agreement with some of the previously known facts.

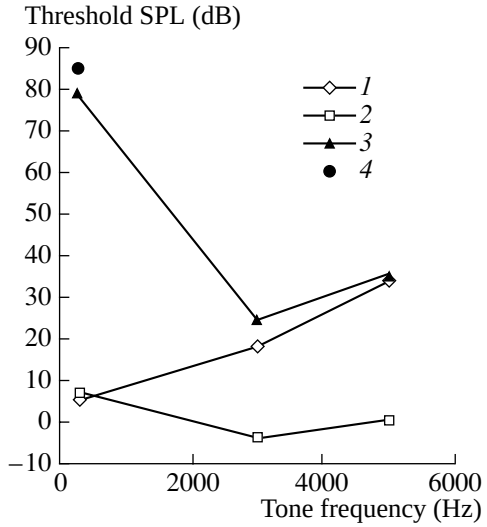


Fig. 7. Absolute hearing thresholds and tone-by-tone masking thresholds (the TVN listener): (*1*) absolute hearing thresholds of the listener, (*2*) absolute hearing thresholds by Zwicker and Feldtkeller, (*3*) masking thresholds of the listener for masking by a 300-Hz tone; and (*4*) the SPL and the frequency of the tone masker (85 dB and 300 Hz).

The values obtained for the BMLD of a 3000-Hz tone in narrowband noise only slightly exceed those typical of high-frequency tones in a broadband masker. They are somewhat lower than the data reported in [2] but correlate fairly well with the results reported in subsequent publications [5–7].

The large values of the BMLD under the half-wave AM conditions are presumably related to the large peak-factor of the envelope: under these conditions, the detection of interaurally antiphase modulation occurs with a considerably enhanced subjective feature of the auditory pattern separation into central and lateral parts. This corresponds to the manifestation of the normalized correlation mentioned in [2–5], because the normalized correlation depends on the peak-factor (unlike the correlation coefficient used more often). The authors of [6, 7], present data demonstrating an increase in the BMLD of high-frequency tones with an increase in the fluctuations of the envelope of a narrowband masker under conventional experimental conditions. The authors compared the detection thresholds of high-frequency tones masked with weakly fluctuating (low-noise) noise and strongly fluctuating (Gaussian, high-noise) noise. However, in the latter case, the BMLD did not exceed 8–10 dB as well.

Presumably, the increase in the BMLD with increasing carrier frequency can be explained by the increase in the critical bandwidth, which covers an increasingly greater part of the spectrum of the half-wave AM (Fig. 3).

The effect of the masking of the low-frequency band of the basilar membrane on the detection thresholds of an interaurally antiphase periodic component in the modulating noise function and on the lateralization

thresholds based on the interaural phase difference of the amplitude modulation of high-frequency signals seems to testify in favor of the nonlinear mechanism of binaural unmasking at high frequencies. However, this hypothesis is difficult to accept. As was noted above, an interaurally antiphase tone under the half-wave AM conditions is detected by the feature of a subjective separation of the right and left lateral parts from the central acoustic pattern (which may also be called the detection of decorrelation). According to the reports of the listeners, they hear and detect this feature not in the low-frequency but in the high-frequency region, where, as we have seen before, the masking of pure tones by the low-frequency masker is practically absent. One can assume that this feature, providing a higher sensitivity in detecting the desired signal, manifests itself in the same way with respect to noise. Finally, it is possible that the masking occurs not on the basilar membrane but higher, although this seems to be unlikely. Note that, under the conventional conditions of masking of a low-frequency tone by a narrowband masker, the half-wave AM conditions are realized spontaneously, because of the detection at reception, but when an interaurally antiphase tone is added to a dichotic narrowband masker, the feature of the separation of the acoustic pattern into three parts does not manifest itself or is barely noticeable, and the BMLD does not exceed 15 dB [10].

Some of our results noticeably differ from the literature data. In particular, we obtained an order of magnitude smaller values for the lateralization thresholds based on the interaural delay of the amplitude envelope (40–80 μs), independently of the masking of the low-frequency band of the basilar membrane [15]. However, one of the recent publications [16] reported on results similar to ours. The authors of this publication measured the lateralization thresholds based on the interaural delay of a 100% sine AM (lateralization thresholds of 200 μs) and a half-wave AM called by the authors "transported tone" (a threshold of 100 μs). The envelope of the latter was invented by the authors to provide the high-frequency channels by the information available in the low-frequency channels, i.e., for the same purpose as in our studies [8–12]. In our experiments, we also observed a decrease in the lateralization thresholds under the half-wave AM conditions.

It is quite possible that some of the discrepancies between our data and the data of other authors are explained by the small number of listeners used for testing and by their age limits.

The main results of our experimental studies are as follows:

—In the case of detecting the periodic component in the amplitude envelope of high-frequency tones, the BMLD is 15–25 dB if, before the modulation, the modulating function is subjected to a half-wave rectification (the half-wave AM conditions).

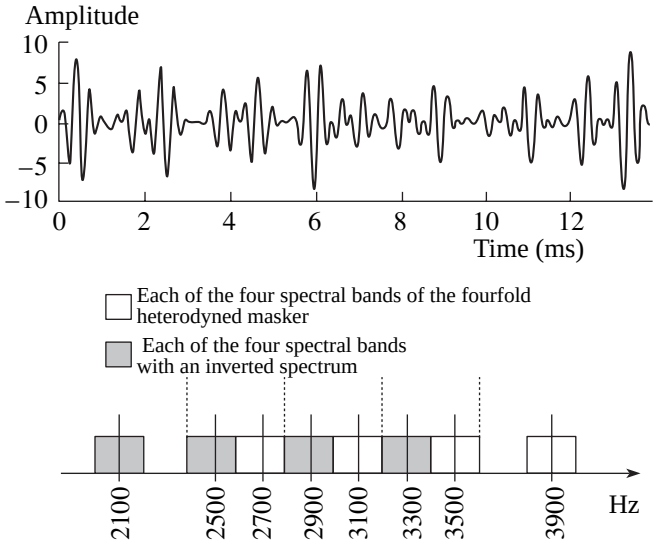


Fig. 8. Time realization and the spectrum of a broadband signal. The dotted vertical lines show the frequencies of the tone components that are present in the polyharmonic carrier (2400, 2800, 3200, and 3600 Hz), which performs multiple heterodyning of the masker, but are absent in the spectrum of the broadband signal. The solid vertical lines show the tone components of the detected polyharmonic signal.

—Under the linear AM conditions, i.e., without the half-wave rectification of the modulating function, the BMLD is within 10 dB.

—Under the half-wave AM conditions, the BMLD increases with an increase in the carrier frequency.

—Under the half-wave AM conditions, the BMLD decreases without any subsequent growth with increasing carrier frequency if the modulating function, after passing through a linear detector, passes through a low-frequency bandpass filter with a cutoff frequency close to the upper frequency of the narrowband masker in the envelope.

—Under the half-wave AM conditions, the BMLD sharply decreases in the case of the masking of the low-frequency band of the basilar membrane mainly at the expense of an increase in the masking threshold of the binaurally antiphase rhythmic component of the modulating function.

—The lateralization thresholds based on interaural phase differences in the low-frequency amplitude envelope drastically increase in the case of the masking of the low-frequency band of the basilar membrane.

—The lateralization thresholds based on interaural amplitude differences do not depend on the masking of the low-frequency band of the basilar membrane.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 03-04-48746.

REFERENCES

1. I. J. Hirsh, J. Acoust. Soc. Am. **20**, 536 (1948).
2. L. R. Bernstein and C. Trahiotis, J. Acoust. Soc. Am. **91**, 306 (1992).
3. L. R. Bernstein and C. Trahiotis, J. Acoust. Soc. Am. **100**, 1754 (1996).
4. L. R. Bernstein and C. Trahiotis, J. Acoust. Soc. Am. **100**, 3774 (1996).
5. L. R. Bernstein, S. Par, and C. Trahiotis, J. Acoust. Soc. Am. **106**, 870 (1999).
6. D. A. Eddins and L. E. Barber, J. Acoust. Soc. Am. **103**, 2578 (1998).
7. J. W. Hall, J. H. Grose, and W. M. Hartman, J. Acoust. Soc. Am. **103**, 2573 (1998).
8. V. N. Telepnev, M. V. Tarasova, and N. A. Dubrovskiĭ, in *Proceedings of X Session of the Russian Acoustical Society* (GEOS, Moscow, 2000), Vol. 2, p. 270.
9. V. N. Telepnev, M. V. Tarasova, and N. A. Dubrovsky, http://www.akin.ru/Docs/Rao/Ses10/sp8.PDF.
10. V. N. Telepnev, M. V. Tarasova, N. A. Dubrovsky, *et al.*, Sens. Sist. **17** (2), 158 (2003).
11. V. N. Telepnev and N. A. Dubrovsky, in *Proceedings of XIII Session of the Russian Acoustical Society* (Moscow, 2003), Vol. 3, p. 247.
12. V. N. Telepnev and N. A. Dubrovsky, http://www.akin.ru/Docs/Rao/Ses13/M23.PDF.
13. J. L. Flanagan and B. J. Watson, J. Acoust. Soc. Am. **40**, 456 (1966).
14. E. Zwicker and R. Feldtkeller, *The Ear as a Communication Receiver* (Acoust. Soc. Am., Woodbury, N.Y., 1998; Svyaz', Moscow, 1971).
15. G. B. Henning, J. Acoust. Soc. Am. **55**, 84 (1974).
16. L. R. Bernstein and C. Trahiotis, J. Acoust. Soc. Am. **112**, 1026 (2002).

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 359–360. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 429–430.

CHRONICLE

Nikolaĭ Andreevich Dubrovsky (On His 70th Birthday)

April 25, 2003, marked the 70th birthday of the prominent Russian scientist-acoustician Nikolai Andreevich Dubrovsky, doctor of physics and mathematics, professor, member of the Academy of Natural Sciences of Russian Federation, director of the Andreev Acoustics Institute, and president of the Russian Acoustical Society.

Dubrovsky's career in science began at the Acoustics Institute in 1957, immediately after his graduation from the Moscow Institute of Physics and Technology, where he specialized in radio engineering and acoustics. Soon after that, the founder of the Acoustics Institute, Academician N.N. Andreev, invited the capable young scientist to work at his department in a new area of research, which included psychological-engineering, biophysical, and psychoacoustic studies.

In 1962, Dubrovsky became head of the Laboratory of Psychoacoustics and Biophysics of Hearing. Fundamental research carried out by Dubrovsky was based on the use of most-advanced physical and mathematical methods. In 1963, he defended his candidate dissertation, which generalized the studies of visual perception and included an analysis of the subjective criteria used by an operator in detecting complex images. Immediately after receiving his candidate degree, Dubrovsky, as a promising young scientist, was offered a visiting fellowship in the United States. There, he spent one whole year working at the leading scientific centers. Upon returning to Moscow, he again headed his laboratory and continued his biophysical studies in acoustics.

Fundamental research into the properties of visual and auditory analysis in both humans and animals was carried out by Dubrovsky in parallel with such well known Russian scientists as V.S. Grigor'ev, G.V. Gershuni, and L.A. Chistovich. Together, these studies formed a new field of research called bionics. The results of these studies were later used in developing fundamentally new methods of signal processing and efficient cybernetic control systems.

Progressively organizing and developing research in bionics at the Acoustics Institute, Dubrovsky became the leader of a group of researchers who carried out comprehensive studies in this area. Under his supervision and with his participation, physiological studies of hearing were combined with the development of morphological studies of neural auditory brain structures and with fine psychoacoustic studies of auditory analysis.

In studying auditory perception, Dubrovsky obtained fundamental results, including the discovery of the critical band and the nonlinear phenomena in the so-called modulation hearing. In addition, important results were obtained in developing models of attention in application to the spatial hearing function and in developing the computer models of auditory periphery.

Simultaneously, Dubrovsky initiated the development of methods for the practical realization of fundamental results in solving new problems and in finding new ways for designing more efficient cybernetic systems. As a result, new instruments were developed for acoustics and underwater acoustics.

In 1968, Dubrovsky organized studies of echolocation and signal processing by the auditory system of dolphins. The most important finding (which later was confirmed by other researchers) was the existence of two auditory subsystems in dolphins: the "active" subsystem used for analyzing echo signals and the "passive" one used for analyzing extraneous sounds. In the active subsystem, Dubrovsky discovered a critical time interval, within which specific mechanisms of acoustic data processing were effective. The results of studying

1063-7710/04/5003-0359$26.00 © 2004 MAIK "Nauka/Interperiodica"

the mechanisms of data processing by the auditory system of dolphins were generalized in the monograph *Sensory Basis of Orientation in Cetaceans* written by Dubrovsky together with V.M. Bel'kovich. This book received wide recognition.

In 1976, Dubrovsky began supervising the projects concerned with cybernetic methods in acoustics and underwater acoustics. The comprehensive studies supervised by him ranged from morphological and functional methods of signal processing by auditory neurons of the brain with auditory image discrimination to the development of modern technical systems based on cybernetic principles of acoustic pattern recognition and to the design of an artificial intellect.

In 1980, Dubrovsky defended his doctoral dissertation, where he generalized the results of his long-term biophysical, bioacoustic, and bionic studies carried out with the use of the modern physical and mathematical approaches. In recent years, Dubrovsky was involved in research related to the acoustic monitoring of the ocean with the aim to study the climate variability and the formation of low-frequency (including infrasonic) acoustic fields produced in the ocean by thunderstorms in the atmosphere. The results of these studies gained wide recognition and were presented at many Russian and international scientific conferences and meetings.

Today, Professor Dubrovsky is deeply involved in educating young scientists. He heads the postgraduate and doctoral departments at the Acoustics Institute. He chairs two dissertation councils of the Supreme Certifying Commission of the Russian Federation. Since 1990, he has volunteered as the Chair of the Department of Acoustic Information Systems of the Moscow State Institute of Radio Engineering, Electronics, and Automation (Technical University). From 1983 to 1996, Dubrovsky delivered lectures on general acoustics at the Department of Hydrocosmos Physics of the Moscow Institute of Physics and Technology. Since 1996, he has chaired the Department of General and Applied Acoustics of this institute. He takes an active part in the process of integration of science and higher education. By now, 2 doctoral and 21 candidate dissertations and 35 graduate projects (carried out by students from different universities of Russia) have been prepared under his supervision.

Dubrovsky is the author of 19 inventions and patents, most of which had been successfully implemented. He has published more than 200 scientific works, including 6 monographs. Scientific results obtained by Dubrovsky and his students have found wide acceptance among the international scientific community.

Dubrovsky is also deeply involved in scientific-organizational and social activities. Since 1989, he has been deputy director of the Acoustics Institute and, at the same time, chaired the Council of the Working Body of this institute. In 1990, Dubrovsky was elected director of the Andreev Acoustics Institute. He is the president of the Russian Acoustical Society, a member of the board of the International Commission on Acoustics, a member of the editorial council of the *Sensory Systems* journal of the Russian Academy of Sciences, a member of the editorial board of the *Acoustical Physics* journal of the Russian Academy of Sciences, a member of the Scientific-Engineering Council of the Russian Ship-Building Agency, etc.

For his long-term fruitful scientific, pedagogical, and organizational activities, Dubrovsky was awarded the Badge of Honor and several medals. He received the title of Honorary Ship Builder from the Russian Ship-Building Agency. He is winner of the USSR State Award and the Prize of the Ministry of Defense Industry of Russian Federation.

In 2003, in addition to his birthday, Dubrovsky celebrated 46 years of his continuous work at the Acoustics Institute. We wish Nikolai Andreevich Dubrovsky new achievements in his scientific, social, and pedagogical activities.

Translated by E. Golyamina

Acoustical Physics, Vol. 50, No. 3, 2004, pp. 361–362. Translated from Akusticheskiĭ Zhurnal, Vol. 50, No. 3, 2004, pp. 431–432.

CHRONICLE

Viktor Vasil'evich Tyutekin (On His 75th Birthday)

On December 3, 2003, Viktor Vasil'evich Tyutekin—doctor of physics and mathematics, professor, head of laboratory of the Andreev Acoustics Institute, and a well-known specialist in acoustics—turned seventy five.

In 1952, Tyutekin graduated from the Radiophysical Faculty of Gorki State University. Then, he became a postgraduate student of the Physical Institute of the Academy of Sciences of the USSR and, in 1955, he received the degree of candidate of science. Among his first teachers, he was lucky to have such outstanding scientists as L.M. Brekhovskikh, Yu.M. Sukharevskii, and G.D. Malyuzhinets. They helped Tyutekin to develop his talent for both comprehensive theoretical studies and fine physical experiments. An important contribution to his education was made by his collaboration with M.A. Isakovich and N.S. Ageeva.

From 1955 to this day, Tyutekin has worked at the Andreev Acoustics Institute. He progressed from junior researcher to head of laboratory (1972). One of his first research projects (his candidate dissertation) was concerned with theoretical and experimental studies aimed at constructing artificial acoustic media. Later, these studies resulted in the development of special hydroacoustic coatings, which combined sound-absorbing, sound-insulating, and vibration-damping properties and found wide industrial application. The results of these studies were generalized in his doctoral dissertation (1968).

Other numerous (more than 200) publications by Tyutekin were devoted to different branches of acoustics. The most important works are briefly described below. In theoretical acoustics, they include the development of methods for solving the problem of the properties of elastic media described by higher-order differential equations. Tyutekin developed a method based on the use of the Riccati tensor differential equation for the so-called elastic impedances. This method allowed the determination of the wave properties of homogeneous and layered-inhomogeneous elastic bodies. In particular, Tyutekin calculated the acoustic characteristics of radially inhomogeneous cylindrical waveguides (in cooperation with E.V. Golubeva), solved the problem of sound wave diffraction by closed shells of revolution (together with V.Yu. Prikhod'ko), performed a synthesis and an experimental study of distributed absorbers for elastic waves in bars and plates (together with A.P. Shkvarnikov), determined the characteristics of elastic waveguides with rectangular cross sections (together with A.E. Vovk), etc.

In cooperation with M.A. Isakovich and V.I. Kashina, Tyutekin developed the principle of the so-called waveguide insulation, which allowed the design of efficient devices for sound and vibration control. On the basis of the theory of microinhomogeneous media (which was developed by I.A. Ratinskaya (Chaban)), Tyutekin (with the participation of R.N. Viktorova and T.B. Golikova) developed sound-absorbing materials on the basis of rubber-like media with heavy inclusions and brought them to the stage of mass production. The efficiency of these materials is practically independent of external static pressure.

Using the theory of active systems for suppressing spatial acoustic fields (which was developed by G.D. Malyuzhinets and M.V. Fedoryuk), Tyutekin designed and experimentally tested models of such systems for single-mode and multimode waveguides (together with D.V. Stepanov and A.A. Mazanikov) and

1063-7710/04/5003-0361$26.00

for a cylindrical waveguide with elastic walls (together with V.N. Merkulov), as well as for spatial active systems of planar (with E.V. Korotaev) and spherical (with V.N. Merkulov) configurations.

At the same time, Tyutekin paid much attention to the development of new measuring techniques in acoustics. Two of them are as follows:

(i) measurement of the characteristics of normal waves in multimode waveguides, including the phase velocities and amplitudes of these waves and the matrices of reflection coefficients for the reflection from obstacles in the waveguide;

(ii) a low-frequency acoustic pipe for measuring the acoustic parameters of materials and structures in the travelling wave mode (the reflection coefficient, the sound insulation, the acoustic impedance, etc.). A distinctive feature of this method is the small wave base achieved with the use of unidirectional sound receivers and an active sound absorber placed at the end of the pipe.

Tyutekin successively combines research and invention. He owns more than 50 inventor's certificates (including one US patent). Half of his inventions are employed in industry and practical research. For his inventor's activity, Tyutekin received the title of Honored Inventor of the Russian Federation (1989). For his tutorial activity (he has educated 2 doctors and 15 candidates of science), he received the title of professor in 1972.

Today, Tyutekin continues working as an expert acoustician combining his high theoretical erudition and scientific productivity with creative enthusiasm. Within the last five years, he has published more than 15 scientific papers devoted to different acoustic problems. In particular, he introduced the definition of a new acoustic object called a multichannel long line and calculated its characteristics. He also solved some problems of synthesizing new sound-absorbing media and carried out theoretical studies aimed at developing active methods of sound suppression on the basis of the field decomposition into spatial harmonics (together with A.I. Boiko).

The editorial board of *Acoustical Physics* heartily congratulates Viktor Vasil'evich Tyutekin (who has been an active member of the board for more than 20 years) on his birthday and wishes him health and further creative work.

Translated by E. Golyamina