200787

AD-200787

NTIS per DDC-IR-10 May 66

UNCLASSIFIED

DATE 3/1/56 NEL NO. 1

~~CONFIDENTIAL~~

AUTH. Lib of Congress ltr of 8/25/54

A M Watson 3520

DATE 12/8/58 NEL NO. 1

RECLASSIFIED TO UNCLASSIFIED

AUTH. Jr. Underwater Acoustics V. 8 No. 8 p. 593 Oct 1958

E Engbring 2420

SIGN.

SUMMARY TECHNICAL REPORT

OF THE

NATIONAL DEFENSE RESEARCH COMMITTEE

Manuscript and illustrations for this volume were prepared for publication by the Summary Reports Group of the Columbia University Division of War Research under contract OEMsr-1131 with the Office of Scientific Research and Development. This volume was printed and bound by the Columbia University Press.

Distribution of the Summary Technical Report of NDRC has been made by the War and Navy Departments. Inquiries concerning the availability and distribution of the Summary Technical Report volumes and microfilmed and other reference material should be addressed to the War Department Library, Room 1A-522, The Pentagon, Washington 25, D. C., or to the Office of Naval Research, Navy Department, Attention: Reports and Documents Section, Washington 25, D. C.

Copy No.

98

This volume, like the seventy others of the Summary Technical Report of NDRC, has been written, edited, and printed under great pressure. Inevitably there are errors which have slipped past Division readers and proofreaders. There may be errors of fact not known at time of printing. The author has not been able to follow through his writing to the final page proof.

Please report errors to:

JOINT RESEARCH AND DEVELOPMENT BOARD
PROGRAMS DIVISION (STR ERRATA)
WASHINGTON 25, D. C.

A master errata sheet will be compiled from these reports and sent to recipients of the volume. Your help will make this book more useful to other readers and will be of great value in preparing any revisions.

SUMMARY TECHNICAL REPORT OF DIVISION 6, NDRC

VOLUME 9

RECOGNITION OF UNDERWATER SOUNDS

OFFICE OF SCIENTIFIC RESEARCH AND DEVELOPMENT
VANNEVAR BUSH, DIRECTOR

NATIONAL DEFENSE RESEARCH COMMITTEE
JAMES B. CONANT, CHAIRMAN

DIVISION 6
JOHN T. TATE, CHIEF

WASHINGTON, D. C., 1946

NATIONAL DEFENSE RESEARCH COMMITTEE

James B. Conant, *Chairman*
Richard C. Tolman, *Vice Chairman*

Roger Adams	Army Representative [1]
Frank B. Jewett	Navy Representative [2]
Karl T. Compton	Commissioner of Patents [3]

Irvin Stewart, *Executive Secretary*

[1] *Army representatives in order of service:*

Maj. Gen. G. V. Strong
Maj. Gen. R. C. Moore
Maj. Gen. C. C. Williams
Brig. Gen. W. A. Wood, Jr.
Col. L. A. Denson
Col. P. R. Faymonville
Brig. Gen. E. A. Regnier
Col. M. M. Irvine
Col. E. A. Routheau

[2] *Navy representatives in order of service:*

Rear Adm. H. G. Bowen
Capt. Lybrand P. Smith
Rear Adm. J. A. Furer
Rear Adm. A. H. Van Keuren
Commodore H. A. Schade

[3] *Commissioners of Patents in order of service:*

Conway P. Coe
Casper W. Ooms

NOTES ON THE ORGANIZATION OF NDRC

The duties of the National Defense Research Committee were (1) to recommend to the Director of OSRD suitable projects and research programs on the instrumentalities of warfare, together with contract facilities for carrying out these projects and programs, and (2) to administer the technical and scientific work of the contracts. More specifically, NDRC functioned by initiating research projects on requests from the Army or the Navy, or on requests from an allied government transmitted through the Liaison Office of OSRD, or on its own considered initiative as a result of the experience of its members. Proposals prepared by the Division, Panel, or Committee for research contracts for performance of the work involved in such projects were first reviewed by NDRC, and if approved, recommended to the Director of OSRD. Upon approval of a proposal by the Director, a contract permitting maximum flexibility of scientific effort was arranged. The business aspects of the contract, including such matters as materials, clearances, vouchers, patents, priorities, legal matters, and administration of patent matters were handled by the Executive Secretary of OSRD.

Originally NDRC administered its work through five divisions, each headed by one of the NDRC members. These were:

Division A — Armor and Ordnance
Division B — Bombs, Fuels, Gases, & Chemical Problems
Division C — Communication and Transportation
Division D — Detection, Controls, and Instruments
Division E — Patents and Inventions

In a reorganization in the fall of 1942, twenty-three administrative divisions, panels, or committees were created, each with a chief selected on the basis of his outstanding work in the particular field. The NDRC members then became a reviewing and advisory group to the Director of OSRD. The final organization was as follows:

Division 1 — Ballistic Research
Division 2 — Effects of Impact and Explosion
Division 3 — Rocket Ordnance
Division 4 — Ordnance Accessories
Division 5 — New Missiles
Division 6 — Sub-Surface Warfare
Division 7 — Fire Control
Division 8 — Explosives
Division 9 — Chemistry
Division 10 — Absorbents and Aerosols
Division 11 — Chemical Engineering
Division 12 — Transportation
Division 13 — Electrical Communication
Division 14 — Radar
Division 15 — Radio Coordination
Division 16 — Optics and Camouflage
Division 17 — Physics
Division 18 — War Metallurgy
Division 19 — Miscellaneous
Applied Mathematics Panel
Applied Psychology Panel
Committee on Propagation
Tropical Deterioration Administrative Committee

NDRC FOREWORD

As events of the years preceding 1940 revealed more and more clearly the seriousness of the world situation, many scientists in this country came to realize the need of organizing scientific research for service in a national emergency. Recommendations which they made to the White House were given careful and sympathetic attention, and as a result the National Defense Research Committee [NDRC] was formed by Executive Order of the President in the summer of 1940. The members of NDRC, appointed by the President, were instructed to supplement the work of the Army and the Navy in the development of the instrumentalities of war. A year later, upon the establishment of the Office of Scientific Research and Development [OSRD], NDRC became one of its units.

The Summary Technical Report of NDRC is a conscientious effort on the part of NDRC to summarize and evaluate its work and to present it in a useful and permanent form. It comprises some seventy volumes broken into groups corresponding to the NDRC Divisions, Panels, and Committees.

The Summary Technical Report of each Division, Panel, or Committee is an integral survey of the work of that group. The first volume of each group's report contains a summary of the report, stating the problems presented and the philosophy of attacking them and summarizing the results of the research, development, and training activities undertaken. Some volumes may be "state of the art" treatises covering subjects to which various research groups have contributed information. Others may contain descriptions of devices developed in the laboratories. A master index of all these divisional, panel, and committee reports which together constitute the Summary Technical Report of NDRC is contained in a separate volume, which also includes the index of a microfilm record of pertinent technical laboratory reports and reference material.

Some of the NDRC-sponsored researches which had been declassified by the end of 1945 were of sufficient popular interest that it was found desirable to report them in the form of monographs, such as the series on radar by Division 14 and the monograph on sampling inspection by the Applied Mathematics Panel. Since the material treated in them is not duplicated in the Summary Technical Report of NDRC, the monographs are an important part of the story of these aspects of NDRC research.

In contrast to the information on radar, which is of widespread interest and much of which is released to the public, the research on subsurface warfare is largely classified and is of general interest to a more restricted group. As a consequence, the report of Division 6 is found almost entirely in its Summary Technical Report, which runs to over twenty volumes. The extent of the work of a Division cannot therefore be judged solely by the number of volumes devoted to it in the Summary Technical Report of NDRC: account must be taken of the monographs and available reports published elsewhere.

Any great cooperative endeavor must stand or fall with the will and integrity of the men engaged in it. This fact held true for NDRC from its inception, and for Division 6 under the leadership of Dr. John T. Tate. To Dr. Tate and the men who worked with him—some as members of Division 6, some as representatives of the Division's contractors—belongs the sincere gratitude of the Nation for a difficult and often dangerous job well done. Their efforts contributed significantly to the outcome of our naval operations during the war and richly deserved the warm response they received from the Navy. In addition, their contributions to the knowledge of the ocean and to the art of oceanographic research will assuredly speed peacetime investigations in this field and bring rich benefits to all mankind.

The Summary Technical Report of Division 6, prepared under the direction of the Division Chief and authorized by him for publication, not only presents the methods and results of widely varied research and development programs but is essentially a record of the unstinted loyal cooperation of able men linked in a common effort to contribute to the defense of their Nation. To them all we extend our deep appreciation.

Vannevar Bush, Director
Office of Scientific Research and Development

J. B. Conant, Chairman
National Defense Research Committee

FOREWORD

METHODS NOT involving the human ear have been developed for the detection and recognition of underwater sounds. These, however, generally supplement rather than supplant aural recognition. In the past the human ear has been most generally used as a detector of underwater sounds, and there is every reason to believe that aural recognition of underwater sounds will continue to be of large importance in subsurface warfare.

The work described in this report was a portion of the research program originally proposed in 1941, which came to include an investigation of all the factors involved not only in the detection of submerged or partially submerged submarines but also in the detection of surface craft. The matter presented is limited to a summary of present knowledge of aural recognition of underwater sounds, and is closely related to three other reports in the Division 6 Summary Report Series; namely, Volume 6A, *Military Oceanography*, by the Woods Hole Oceanographic Institution, Volume 7, *Principles of Underwater Sound*, by Dr. Carl Eckart, and Volume 8, *Physics of Sound in the Sea*, by the staff of the Sonar Analysis Group. The physical listening devices developed by the Division are not described in Volume 9 but may be found in Volume 14 of this series, *Sonar Listening Systems*.

The experimental work of the Division was undertaken principally at its San Diego Laboratory under the contract with the University of California. Studies of experimental results and their application were continued by the Sonar Analysis Group operating until lately under a Division contract with Columbia University.

It was most fortunate that in undertaking an investigation of aural recognition and its application to subsurface warfare the Division had available the results of previous research on performance of the ear. In this connection special mention should be made of the work of Dr. Harvey Fletcher and his associates in the Bell Telephone Laboratories, Drs. S. S. Stevens and H. Davis at Harvard University, as well as others who continued to make substantial contributions to the field of psychoacoustics.

The preparation of this report was undertaken by Dr. A. Spector, under the general supervision of Dr. Lyman Spitzer, Jr., Director of the Sonar Analysis Group of the Bureau of Ships. To them the Division expresses its deep appreciation.

JOHN T. TATE
Chief, Division 6

PREFACE

Sounds carry well in the sea, better than they do in air, and enormously better than do electromagnetic disturbances such as light. Underwater sounds are therefore well adapted to providing information about relatively distant objects and events on or beneath the surface of the sea. Thus, the detection of such sounds has important uses in subsurface warfare.

The detection of sound requires, in addition to various types of receiving and amplifying equipment, a detector which can recognize when a particular sound is present. To date, the human ear has been most generally used as a detector of underwater sounds, and its performance has received the greatest amount of study. This volume summarizes present knowledge concerning the aural recognition of underwater sounds. Information on the factors affecting aural recognition, and ability to predict whether or not a particular sound will be recognized under various conditions are useful in a number of ways, as, for example, in the training and selection of sonar operators, in the design and operation of gear, and in estimation of the maximum effective operating range of the sonar gear in each specific situation.

Research in this field during the war has been primarily practical in nature. Emphasis has been placed on specific answers to questions of operational importance, rather than on a broad understanding of the basic factors involved. For the most effective use of these results, however, and for planning future research in this field, a more basic approach is necessary. For this reason current theory on the performance of the ear, a subject known as *psychoacoustics*, has been introduced into the discussion wherever it seemed profitable. While this theory, developed by Fletcher, Stevens, and others, has not yet reached the point where it permits conclusive predictions of the behavior of the ear, it does indicate a certain coherence among the observations that would otherwise not be apparent. Even where the theory seems inconsistent with the observed results, the theoretical discussion may be profitable in suggesting a new experimental or theoretical approach.

While the basic theory of the performance of the ear has been kept in mind in the preparation of this volume, a deliberate attempt has been made to focus attention also on the many practical aspects of the information discussed here. Thus, in the discussion of the data, many points are brought out which are not directly relevant to results on the performance of the ear, but which bear on the application of these results to the most effective use of sonar in subsurface warfare. It is hoped that the resultant volume may be of interest both to the technical worker in psychoacoustics and to the engineer interested in the application of psychoacoustic data to the design and operation of equipment. While the discussion is relatively exhaustive, the general level of the writing has been made sufficiently elementary so that it may be understood by those with little experience either in mathematics or in sonar research.

In reading this volume several important distinctions should be kept in mind. The most important is that between "wanted" sounds, for which the operator is listening, and "unwanted" sounds, or interference tending to mask the wanted sounds. A wanted sound is frequently called a *signal*, while the many unwanted sounds are grouped together under the heading *background*. Whether a particular sound is classified as signal or background depends on the circumstances. For example, the sounds produced on board a submarine add to the background when the equipment mounted on the submarine is being used for listening to enemy vessels. When these local sounds are being monitored by the submarine, however, they constitute the signal. Similarly, sounds from a submarine are the signal used by listening gear on an antisubmarine vessel, but form part of the background when echoes are used to measure the range to the submarine.

A second, less fundamental distinction is that between local sounds, associated with the sonar gear and its mounting, and more distant sounds, produced by sources far away. Local sounds, including electrical noise in the hydrophone circuits, motion of water around the hydrophone or ship (if the hydrophone is mounted on board ship), or noise produced by machinery on board ship, are frequently grouped together under the heading *self-noise*, and generally constitute much of the unwanted noise background. Sounds from distant objects include many generally unwanted components, such as noise from whitecaps forming on the sea surface, or sound scattered from irregularities on the bottom. In addition, some distant sounds, such as the sound generated or reflected by surface vessels, submarines, torpedoes and the like, usually are the signals which are to be recognized. These sounds may be used to provide information about the bearing and nature of the sound source, and under certain conditions can be used to determine the speed, range, course, change of course, and other maneuvers of the source.

The most important application of sonar hitherto has involved the use of shipboard-mounted gear to detect other ships, either on the surface or submerged. In the discussion of the observations, given in this volume, this situation is usually assumed. It should be emphasized, however, that the basic results have a much wider validity. Sound gear can also be towed, suspended from a buoy, or installed on or under a harbor bottom. The received sounds may be presented to the ear not only on board ship but also in aircraft or at shore stations. While the tactical problems are quite different in each of these cases, the general behavior of the ear, as analyzed in this volume, will be the same.

In Part I, comprising Chapters 1 and 2 of this volume, are presented the relevant facts on the structure and performance of the ear as revealed primarily by published psychoacoustical studies before World War II. Part II, consisting of Chapters 3 through 6, analyzes the data on the masking of sounds produced by a distant target, while Chapters 7 through 11 which constitute Part III deal with the corresponding results for echoes. For convenient practical application, the chief results found in these two parts are briefly stated in Chapters 6 and 11. Most of the material presented in Parts II and III was obtained during the war, in part by the British, and in part by the three following groups under contract with Division 6 of the National Defense Research Committee: the University of California Division of War Research at the U. S. Navy Electronics Laboratory (formerly the U. S. Navy Radio and Sound Laboratory), San Diego, California; Columbia University Division of War Research at the U. S. Navy Underwater Sound Laboratory, New London, Connecticut; and Bell Telephone Laboratories, New York, New York.

This volume has been written by Dr. Aaron Spector, who has introduced much analytical discussion which is not available elsewhere. Final preparation and editing of the material has been carried out cooperatively by the staff of the Sonar Analysis Group. The laboratory groups whose work is reported here have been extremely helpful in communicating results in advance of publication and in discussing the many problems of interest in this subject. Particular acknowledgement is due to the University of California Division of War Research, and to Bell Telephone Laboratories.

LYMAN SPITZER, JR.

CONTENTS

Chapter 1

STRUCTURE OF THE EAR

FUNDAMENTAL TO ANY discussion of the ear and hearing are references 1 through 5. These works, and the literature cited in them, are the sources for all statements about the structure and behavior of the ear which are made in this report and which are not specifically attributed to some other source. A few extensions and applications of the fundamental concepts have been prepared especially for the present report. These are relatively minor and are not identified. Although an effort has been made to specify the limitations of present knowledge, it is clearly impossible to discuss all the details in a condensed summary of this kind. Careful reading of the original publications is necessary for a thorough understanding of this subject.

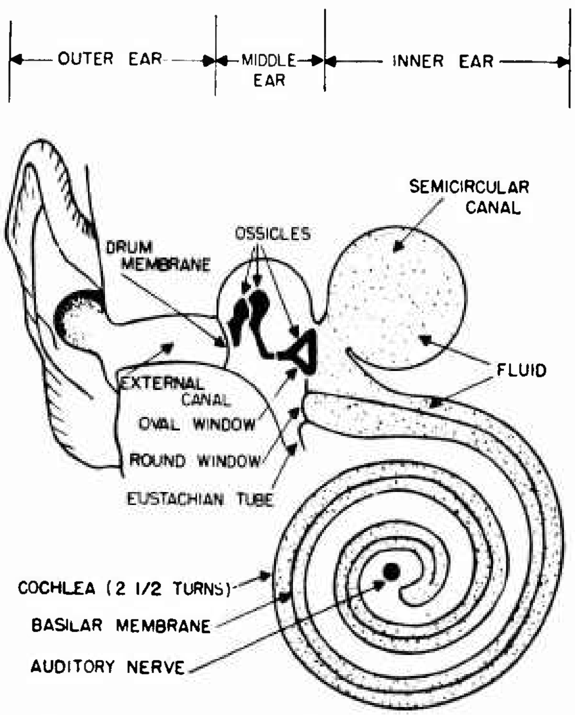


FIGURE 1. Schematic diagram of the ear.

A few standard results of physical acoustics are assumed without proof in this volume. For a discussion of these, the reader is referred to references 6 and 7.

The potentialities of the ear as a detector depend on its nature as a physical mechanism. A brief account of present knowledge about the structure and behavior of the ear is given in the following two chapters and the concepts developed there are applied in the remaining chapters of this volume to the data obtained in studies of sound gear.

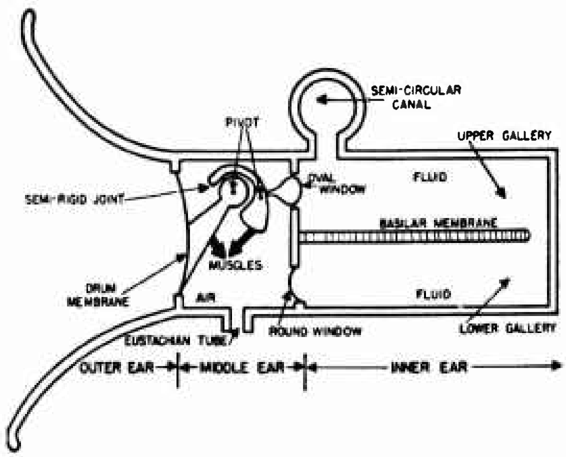


FIGURE 2. Schematic diagram of the outer, middle, and inner ear.

The ear structures are minute and delicate. Few of them are accessible to direct observation when the ear is intact.[8] Most of our ideas about the hearing process are consequently based on less direct lines of evidence. Some of the more important features of the normal human ear are illustrated schematically in Figures 1 through 8. These structures are believed to behave in the following way during the act of hearing. Audible sound waves which enter the external ear and pass through the *ear canal* strike the *drum membrane* (see Figure 1). The resulting motion of this membrane is transmitted by the *ossicles*, a flexibly jointed chain of three small bones stretching across the *middle ear*, to the *inner ear* which is the essential organ of hearing.[a] The nerve impulses generated in

[a] Destruction of the outer ear, or of the ossicles, interferes with hearing; destruction of the inner ear or of the auditory nerve produces total deafness.

the inner ear move along the *auditory nerve* to the brain, where they produce the sensations of hearing.

Present evidence indicates that perceptions of pitch which are aroused by stimulating the ear with tones of different frequency are correlated primarily with *positions* of stimulation pattern on the sensory membranes of the inner ear. This point of view is generally known as the *place theory*. Its adequacy for frequencies below 100 cycles has not been established. The fact that a particular place on the sensory membrane is stimulated by a given tone is considered to result from the selective tuning of the membrane, so that different tonal frequencies produce vibration patterns on this membrane which have their maximum amplitudes at different places. The latter explanation of the ear's behavior is called the *resonance theory*. The place theory and the resonance theory are supplementary; taken together, they provide a consistent understanding of the bulk of the known facts. Continued study may require modification of these points of view, which are nonetheless useful at present in organizing the data and in suggesting further experimental work. The significance and the application of these theories can best be understood from an examination of the data regarding the inner ear structures; these data are presented in the following section.

1.1 INNER EAR

The inner ear structures are contained in a coiled bony tube, named the *cochlea* from its

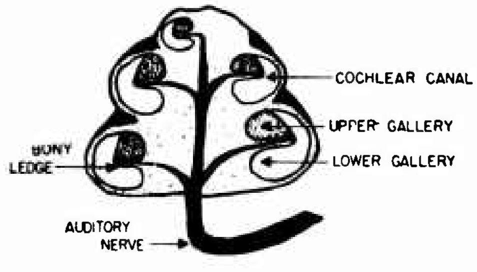


FIGURE 3. Section through the cochlea.

resemblance to a small snail shell. To shield the inner ear from injury and to insulate it partly from the effects of sounds which are not directed to enter it through the external ear canal, the cochlea is enclosed within the wall of the skull. The compact form of the cochlea probably results in greater mechanical strength.

A sectional view of the cochlea is shown in Figure 3. Figures 4 and 5 show how it would look if unrolled. It is some 31 millimeters in length, wider at its base than at its apex, partitioned along its length by the sensory mem-

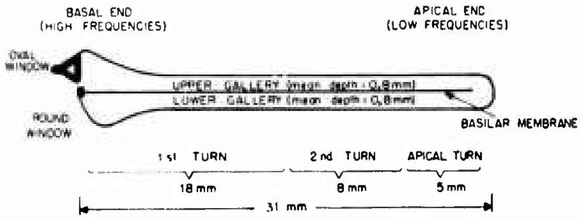


FIGURE 4. Sectional view of the basilar membrane, oval window, and the round window.

branes, and completely filled with clear watery fluid. The walls of the cochlear duct are rigid except for two openings at its base called the *oval window* and the *round window*, from their shapes. These windows adjoin each other on the surface of the middle ear cavity. The round window is sealed by a thin elastic membrane. The oval window has one end of the chain of ossicles secured to its margins by means of tis-

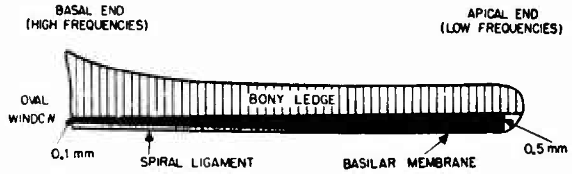


FIGURE 5. Sectional view of the basilar membrane, bony ledge, and the spiral ligament.

sues, in such a manner that motion of the ossicles is relatively free, although the inner ear fluid cannot escape; the other end of the chain of ossicles is attached to the drum membrane. Increase of pressure at the drum displaces it and the ossicles inward, shifting the inner ear fluid and causing the membrane at the round window to bulge. This motion of the cochlear fluid stimulates the auditory nerve.

As shown in Figures 1 and 2, the cochlea is part of a closed system of communicating tubes, the so-called labyrinth, which contains several sense organs and is filled with liquid. Since the walls of the labyrinth are relatively rigid, except for the round and oval windows, motion of the middle ear structures causes mass displacement of fluid within the cochlea only. The structures of the labyrinth other than the cochlea are not described here because they are concerned with bodily equilibrium and not with hearing.

The cochlear canal resembles a spiral staircase wound around a bony central shaft (see Figure 3). This canal is incompletely divided along its length by a bony ledge; the partition of the canal into an upper and a lower gallery is completed by a thin membrane (the *basilar membrane*) in which are embedded throughout its length a great number of fine fibers (the *auditory strings*). The auditory strings lie parallel to the diameter of the cochlear canal and are maintained in tension by the *spiral ligament*. The basilar membrane does not extend quite to the apex of the cochlea (see Figures 4 and 5); the edges of the membrane, with the exception of its apical edge, are fastened to the walls of the cochlear duct. Thus, fluid may pass from one gallery into the other only through the gap at the apex of the basilar membrane.

Since the round and oval windows lie on opposite sides of the basilar membrane, the piston motion of the ossicle at the oval window will produce displacement of the cochlear fluid, either by causing it to flow through the gap at the extreme end of the basilar membrane, or by moving it along a hydraulic short circuit

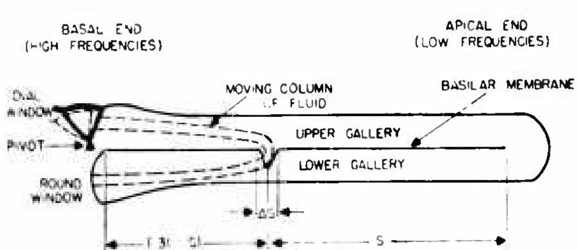


FIGURE 6. Assumed motion of the inner ear structures in response to sound.

(illustrated in Figure 6) which is accompanied by bulging of the basilar membrane. The actual hydraulic path, according to the resonance theory, depends upon the impressed frequency.

1.2 RESONANCE THEORY OF HEARING

It is assumed that the basilar membrane is selectively tuned in the manner of a reed-type frequency meter so that different portions of the membrane will vibrate in synchronism with those impressed frequencies to which they resonate. The possibility that sensations of pitch are associated with the wavelengths of the received sounds may be excluded, since the dimensions of the cochlea are so small compared with the wavelengths of audible sounds. Also, observation indicates that the auditory nerve and brain do not directly analyze the frequency of a stimulus. For example, direct stimulation of the auditory nerve with alternating current of any frequency invariably produces the sensation of pitchless noise.

In the present section, the physical characteristics of the resonant mechanisms of the inner ear will be examined. Assuming that the dynamical behavior of the basilar membrane may be satisfactorily approximated by considering it as an assembly of loosely coupled strings, an arrangement resembling a Venetian blind, it follows that the region of the membrane which responds best to a particular frequency is determined by the combined effect of several factors: the lengths of the auditory strings l; their tensions T; and their mass per unit length m, including the effective increase of mass due to the column of fluid which must be moved when some portion of the membrane moves.

In the following paragraphs the dependence of the resonant frequency on these physical properties of the inner ear is examined. Although the available data do not yield precise values of the relevant quantities, the investigation indicates that resonance of the basilar membrane for perceived sonic frequencies is not unreasonable physically. Moreover, this analysis yields a useful if approximate expression for the position of maximum stimulation on the membrane, for sound of a particular frequency.

The basic expression for the resonant frequency f of a string under a tension T dynes, of length l, and with a mass m per unit length is

$$f = \frac{1}{2l}\sqrt{\frac{T}{m}}\,. \tag{1}$$

This ignores the effects of damping, which would diminish the value of the resonance frequency somewhat. Such damping would be expected from the effects of friction and viscosity in the various parts of the ear. Damping, together with the fact that adjoining auditory strings are coupled by the ground substance of the basilar membrane, should result in a broadened, rather than a sharply defined, response; there is evidence that this occurs, particularly for large amplitudes (see Figures 2 and 3 of Chapter 2). The advantage of damping is that sensations of sound are extinguished soon after the stimulus is removed.

The relation between the position s of a string on the basilar membrane and the frequency to which it responds may be estimated from the mechanical constants of the inner ear (see Figures 4, 5, and 6). It seems reasonable to infer from these constants that the apical end of the membrane is tuned to low frequencies; and the basal end, to high. The auditory strings at the apical end, for example, are approximately five times as long as those at the basal end; the width of the membrane decreasing fairly uniformly from about 0.50 millimeter at the apex to about 0.10 millimeter at the base. The length, in centimeters, of an auditory string may therefore be represented as follows:

$$l = (5.0 \times 10^{-2}) - (1.3 \times 10^{-2})\,s, \tag{2}$$

where s is the distance of the string from the apex of the basilar membrane in centimeters; thus, at the base of the membrane, s equals 3.1 centimeters.

The tension exerted by the spiral ligament on the shorter basal strings is believed to exceed that exerted on the apical strings since the ligament is not a uniform structure. It consists of a few frail fibers at the apical end and thickens gradually toward the basal end of the membrane. The simplest relation to try, in the absence of more definite information, is one which makes the tension increase linearly with distance from the apex, that is,

$$T = as, \tag{3}$$

where s is subject to the same restriction as before and a is a constant. Equation (3) is, of course, somewhat arbitrary; lack of information concerning T is perhaps the greatest gap in our knowledge of the physical parameters affecting the resonance of the inner ear.

The effective mass of an auditory string depends on the liquid loading and may be taken as proportional to the mass of the moving column of fluid. The mass of piano strings is increased in similar fashion, without adding significantly to their rigidity, by coiling copper wire on them. As indicated in Figure 6, the length of such liquid columns is very nearly $2(3.1 - s)$, and the expression for the mass per unit length of a string may therefore be written

$$m = b\,(3.1 - s), \tag{4}$$

where b is assumed here to be a constant, depending on the density of the cochlear fluid. As shown in the next section, there is some reason to believe that b may vary somewhat along the membrane. In view of the lack of specific information about T, this complication is not considered here.

If the motion of the ossicle at the oval window is sufficiently slow, pressure above and below the membrane will remain equal. Hence, it will not be displaced from its equilibrium position along its length, and cochlear fluid will pass from one gallery to the other through the gap at the apex of the membrane, producing simultaneous displacement of the round window. For sufficiently rapid displacements of the ossicle, or for high enough frequencies of vibration, the cochlear fluid has considerable impedance. Hence, when the ossicle thrusts inward, the pressure in the upper gallery will exceed that in the lower, and the membrane will be displaced as shown in Figure 6. Those portions of the membrane with natural frequencies equal, or nearly equal, to the impressed frequency will move with greater amplitudes since their motions can remain in phase with the driving force. While general forced motion of the entire membrane is to be expected, especially for sounds of high intensity, relatively small amplitudes should be found at all points whose natural frequency differs from that of the stimulus tone.

By substituting equations (2), (3), and (4) into equation (1), and denoting the fractional distance from the apex to the base of the membrane, that is, $s/3.1$, by the letter x, we obtain

$$f = \frac{10\sqrt{a/b}}{1 - 0.8x}\sqrt{\frac{x}{1 - x}} = \frac{1300}{1 - 0.8x}\sqrt{\frac{x}{1 - x}}. \tag{5}$$

The constant $\sqrt{a/b}$ has the dimensions centimeters per second. A numerical value of 1,300 has been assigned to this quantity to give the best agreement with the observed data (see Figures 9, 10, and 17 in Chapter 2). Since equation (5) has been derived from several quite approximate assumptions, it cannot be regarded as rigorous. However, it should give results correct to within an order of magnitude. Comparison with the data in Section 2.2 and 2.3 shows that this equation has, in fact, the proper form. It may therefore be regarded as a semi-empirical relationship which provides a useful summary of a large number of observations. For example, when x increases by equal increments, f grows by successively larger amounts. In other words, the positions on the basilar membrane corresponding to two frequencies which differ by a constant number of cycles are closer together for high frequencies than for low. Furthermore, when x is 0.5, f equals 2,200 cycles per second, that is, the center of the basilar membrane is tuned to a frequency a little over 2 kilocycles. Both these conclusions are in agreement with observation (see Section 2.2.1).

The resonance theory of hearing thus provides a correlation, given by equation (5), between the position of an auditory string and its resonance frequency. This prediction is in good agreement with observation. However, if this theory of the ear's structure is to be regarded as creditable, it must provide acceptable answers to several other questions.

For example, are the estimated tensions in the auditory strings of reasonable magnitudes? Is the required rate of change of tension between neighboring strings immoderately large? If certain simplifying assumptions are made, orders of magnitude may be computed for these two quantities. The computations are given in the following subsection.

1.2.1 Tension in Basilar Membrane

Solving equation (1), we obtain

$$T = 4f^2l^2m. \tag{6}$$

If the required tension is unduly large, this fact would show up most clearly at the higher frequencies. It will be assumed, therefore, that f equals 10 kilocycles, which is within an octave or so of the highest audible frequencies. From equation (5) it follows that x equals 0.85 for a frequency of about 10 kilocycles; hence s, the position of the string, is about 2.6 centimeters from the apex of the membrane, or about 0.5 centimeter from the oval window. The length of a string at this position is, from equation (2), approximately 1.6×10^{-2} centimeter.

In addition, m must be evaluated; this quantity is the mass M of the liquid column loading the string, divided by the length l of the string. The evaluation of m is not a simple matter but depends on the solution of the hydrodynamical equations, taking into account the physical properties both of the cochlear canal and of the basilar membrane. An upper limit may be derived by assuming that all the fluid in the cochlear canal between the oval window and vibrating section of the membrane vibrates to and fro. This neglects the effect of viscosity and also neglects the possibility that at the higher frequencies the wavelength of sound in the cochlear canal may be less than the total length of the canal. On this simple assumption, the total volume of the oscillating fluid is $2A(3.1 - s)$, where A is the cross section of each of the two canals, and $(3.1 - s)$ is the length in centimeters of the vibrating column. The factor 2 takes into account the fact that the liquid in both canals takes part in the oscillation. To find the total mass, this volume must be multiplied by the density ρ of the cochlear fluid. This density is 1.034 grams per cubic centimeter at body temperature and may be set equal to unity for these calculations.

This oscillating mass has a cross-sectional area A which is considerably greater than the area $l\Delta s$ of the vibrating region of the membrane (see Figure 6). As a result, the maximum fluid velocity is less than the maximum velocity of the membrane by a factor $l\Delta s/A$. Since the mass is important through its contribution to the kinetic energy, the effective mass is roughly equal to the oscillating fluid mass multiplied by $(l\Delta s/A)^2$. This rather simple approach is, of course, not rigorous but gives the same general result as a more detailed analysis of the potential and kinetic energies of the system.

To find the mass loading per string, the total vibrating mass must be divided by the number of strings which are set vibrating by a pure tone. This number is about equal to 100 for a tone of moderate loudness (see "Beats" in Section 2.1.2). Thus we have finally

$$m = \frac{M}{l} = \frac{2.4(3.1 - s)}{100l}\left[\frac{l\Delta s}{A}\right]^2. \qquad (7)$$

If we substitute equation (7) into (6), let A equal 6×10^{-3} square centimeter, and Δs equal 3×10^{-2} centimeter (see "Beats" in Section 2.1.2.), we find at a frequency of 10 kilocycles,

$$T = \frac{8f^2l^3(\Delta s)^2(3.1 - s)}{100A}$$

$$= \frac{8(10^8)(1.6 \times 10^{-2})^3(3 \times 10^{-2})^2(5 \times 10^{-1})}{10^2 \times 6 \times 10^{-3}}$$

$$= 2.4 \text{ dynes}$$

corresponding to a force of about 2.5 milligrams. Since the cross section of a single string in the membrane is about 10^{-7} square centimeter, the total stress is 2.5×10^7 dynes per square centimeter, or about 0.17 ton per square inch.

This result may be compared with the breaking strength of spider thread, 15 tons per square inch, and of silk, 32 tons per square inch. Since equation (7) gives an upper limit for m, the stress found from the resonance theory appears very plausible. Furthermore, the rate at which the estimated tension changes along the length of the membrane, in other words, the quantity a in equation (3), is fairly small, totaling about 1 dyne per millimeter (25 dynes per 26 millimeters).

1.2.2 Hydrodynamic Theory

For the sake of simplicity the discussion here has been built around the most elementary form of the resonance theory. A more general treatment will be found in reference 9. Consideration is given there to the hydrodynamic equations appropriate to the situation in which a pressure pulse is initiated at one end (the oval window) of a liquid-filled tube with three rigid walls and one flexible wall, corresponding to the bony structure of the cochlear duct and the yielding structure of the basilar membrane. The analysis indicates that the pressure gradients associated with successive cycles of a stimulus tone are propagated along such a tube with a speed of about 50 m per sec. This low velocity arises from various viscous and frictional factors and is in fair agreement with the velocity of approximately 20 meters per second derived from the functional response of the ear. Furthermore, this theory indicates that the amplitudes of the traveling wave will be greatest at specific points of the basilar membrane, and that the positions of these points are associated with the frequencies of the impressed tones. The agreement between the predicted and observed positions is inferior to that obtained with equation (5) which may be due to somewhat uncertain knowledge of the constants needed for the development of the analysis and also to the effects of various simplifying assumptions.

The time required for a disturbance, moving at a velocity of 2,000 centimeters per sec, to pass over the distance of 3.1 centimeters between the oval window and the apex of the membrane is about 1.5 milliseconds. Hence, the high-frequency content of a complex acoustic disturbance made up of all frequencies is sensed about 1.5 milliseconds earlier than the low-frequency content.

The end organs of equilibrium are essentially plumb-bobs suspended in the fluid of the semicircular canals. When these are displaced, a sense of disequilibrium is produced. Pressure variations in the cochlear fluid are transmitted to the semicircular canals, since the cochlea is but one of a group of communicating tubes composing the labyrinth. Ordinarily, pressure changes in the cochlea, produced by incident sounds, cause no motion of the equilibrium receptors in the semicircular canals, because the pressures at any point are independent of direction. This ceases to be true in the case of rotational, or vortex, motion; and it has been pointed out[10] that vertigo and reflex motion of the head in response to very loud sounds is probably associated with the production of eddies in the cochlea.

1.3 SENSORY AND NERVOUS STRUCTURES

Thus far we have spoken of the basilar membrane as though it alone were responsible for the perception of tones. Actually, the motion of the membrane which is produced by tonal stimulation causes a localized disturbance of sensory cells attached to the upper surface of the membrane, and this disturbance of a particular group of sensory cells causes specific fibers in the auditory nerve to respond.

The arrangement of the sensory structures supported by the basilar membrane is shown in Figures 7 and 8. A sequence of about six thousand adjoining arches projects from the surface of the basilar membrane into the upper gallery (Figure 4). These are called *Corti arches* after their discoverer. They are uniformly spaced over the entire length of the basilar membrane. Each of the arches is formed from two stiff bristles which are joined

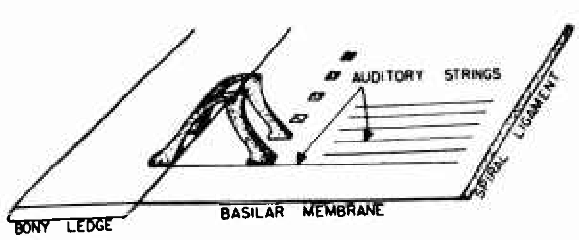


FIGURE 7. View of the basilar membrane through the surface of the upper gallery, showing the Corti arches.

at their upper ends to form a fairly rigid truss. One leg of each arch is supported on the margin of the bony ledge; the other rests on the basilar membrane. The base of an arch measures about a quarter the width of the membrane; in other words, the arches increase in size along with the auditory strings. Thus, as the auditory strings vibrate, the arches rock back and forth, pivoting upon the leg which rests on the bony ledge. The exact function

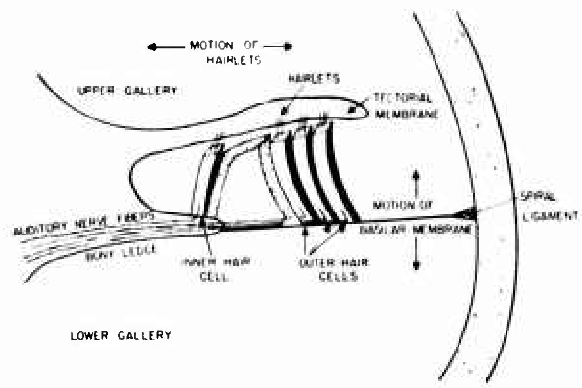


FIGURE 8. View of the basilar membrane showing the Corti arches and hair cells.

which the Corti arches play in the mechanism of hearing is not certain.

As shown in Figure 8, each of the arches is flanked by four sensory cells. These cells form two rows running the length of the basilar membrane. The outer row, which rests on the basilar membrane, is three cells deep; the inner row, which rests on the bony ledge, is only one cell deep. These sensory cells are called *hair cells* because of the tiny hairs which project from their upper surfaces. The hair cells are the end organs of hearing; that is, the fibers of the auditory nerve terminate in the hair cells (see Figures 3 and 8).

Thus, the auditory nerve resembles a telephone cable supplying a distribution board with about 25,000 terminals (6,000 arches with four hair cells per arch). The nerve trunk enters the cochlea through the bony shaft around which the cochlear duct is coiled. Its individual fibers branch out to the hair cells through the bony ledge of the cochlear canal (see Figures 3 and 8).

All the hairlets which project from the free ends of the hair cells are embedded, as shown in Figure 8, in a soft structureless membrane (the *tectorial membrane*) lying above the hair cells. One margin of the tectorial membrane is secured to the bony ledge; the other margin is free. The hairlets, being embedded in the tectorial membrane, cannot follow the rocking motion of the hair cells and Corti arches which is produced by vibration of the auditory strings; instead they are bent to a degree which depends upon the amplitude of motion of the basilar membrane. It is believed that this bending of the hairlets and possibly the compressions and elongations which they suffer stimulate the fibers of the auditory nerve. Thus "hearing" may be compared to a highly differentiated sense of touch.

This description of the action of the hair cells is in agreement with several experimental observations. When the row of three outer hair cells is congenitally missing (in animals with hearing apparatus similar to that of man) and the row of inner hair cells is intact, the animal is partly deaf. In such cases it is necessary to raise the level of a tone about 30 to 40 decibels above the normal threshold of audibility before that tone can be heard. This implies that the row of three outer hair cells responds to faint sounds and that the inner hair cells respond only to loud sounds. If the hair cells are thought of as mounted upon a lever with its fulcrum located on the margin of the bony ledge (see Figure 8) it will be observed that the three outer hair cells, which are situated on the long arm of the lever, will undergo a larger displacement than will the inner hair cells, for any given amplitude of motion of the basilar membrane.

The converse observation to that described above is the following. By continued exposure to loud tones, animals with inner ear structures like that of man can be partly, but permanently, deafened to the tones to which they are exposed. The extent of this deafening never exceeds 30 to 40 decibels; that is, the experimental animals can always hear such tones subsequent to exposure if the intensity level is raised the stated amount above normal threshold. In other words, the receptors for the fainter sounds are distinct from those which respond only to loud sounds.

1.4 ELECTRICAL ACTIVITY

Two methods are commonly used to determine the sensitivity of animals to sound. In the first place, they may be conditioned to associate some pleasant or painful experience with a particular sound which invariably accompanies that experience in the experimental situation. After repeated association of this kind, the test animal gives the same observable response, such as flight or salivation, to the sound that it would normally give to the experience with which the sound is connected. By diminishing the intensity of the test sound until no response to sound is obtained, the auditory threshold of the experimental animal can be determined.

The second method is more direct and provides some results which are probably not obtainable through use of the first. It is based on the observation that stimulation of the ear by sound produces a measurable change of electrical potential at the cochlea. These cochlear potentials may be detected by placing one electrode on the surface of the head and the other in contact with the cochlea, after it has been exposed surgically. These changes of potential are of the order of microvolts, and amplification is required for detailed study. The cochlear potentials are strongest when one electrode makes contact with the cochlea but they can be detected almost anywhere on the head if sufficient amplification is used. Wave form, and changes of wave form or intensity, of sound waves which are used to stimulate the ear are very faithfully reproduced by the cochlear potentials. Thus, speech entering the ear of an experimental animal is intelligibly reproduced by an amplifier and loudspeaker circuit which is actuated by the cochlear potentials. Standard indicating instruments for such studies are the cathode-ray screen and the wavemeter. The cochlear potentials obviously furnish a very powerful tool for investigating such effects as the change of auditory acuity which is produced by long exposure to loud tones.

When an alternating current is passed through the head under appropriate conditions, sensations of tone are produced whose pitch corresponds to the frequency of the current. This phenomenon is termed the electrophonic effect. Formally, at least, it is the inverse of the cochlear potentials which are produced by conversion of mechanical energy into electrical.

Several other matters related to the cochlear potentials are worth discussing at this point. To begin with, the cochlear potentials must be distinguished from the potentials associated with the activity of the fibers in the auditory nerve. When a nerve fiber, including the auditory nerve fibers, responds to a stimulus, this action is accompanied by the development of a potential along the fiber which is never observed in the quiescent state of the nerve. Such potentials are therefore called the *action potentials* of the nerve. When the auditory nerve is severed from the cochlea, the action potentials can no longer be initiated by means of sound. However, chemical or mechanical stimulation may be substituted. Under these conditions, action potentials and cochlear potentials continue to be observed independently of each other. In other words, the auditory nerve is not essentially a passive cable which transmits the cochlear potentials to the brain; its role in the hearing process can be understood only in terms of its behavior as a nerve.

Nerve fibers, including the fibers of the auditory nerve, respond to chemical, mechanical, and electrical stimulation. They do so only when the stimulation exceeds a threshold value. This threshold differs from nerve to nerve, it is different for the various fibers of a given nerve, and it depends upon the previous stimulation which a given fiber has received. Furthermore, the strength of the action potential which develops when the response threshold of a nerve is exceeded is independent of stimulus intensity; in other words, a nerve discharge either occurs at full strength or it does not occur at all. Nerve response is therefore "all or none" in character and is frequently compared to the firing of a gun. Nerve discharge resembles gunfire in one other respect; the fiber needs time to "reload," or, more exactly, it enters a refractory phase after discharge and cannot discharge again until it recovers its sensitivity. The duration of this refractory period is about 1 millisecond (10^{-3} second); hence the maximum rate of discharge of any nerve fiber is about one thousand per second. The cochlear potentials, on the other hand, are continuous functions of the stimulus and can follow the latter to very high frequencies.

The origin of the cochlear potentials and the part they play in hearing are uncertain at the present time. However, they mirror the mechanical events occurring in the cochlea accurately and are therefore useful in studying these events. For example, comparison of action with cochlear potentials shows that the nerve fibers which are stimulated by a given tone fire in synchronism and that they are excited during the rarefaction of the sound wave, that is, when the basilar membrane moves up against the tectorial membrane.

1.5 OUTER AND MIDDLE EAR

The foregoing discussion of the structure of the ear has emphasized the inner ear, since this is the organ where different frequencies are distinguished and where mechanical oscillations are converted into nerve impulses. For a better understanding of the auditory process, however, information on the other parts of the ear is also important. The structure and behavior of the outer and middle ear are therefore discussed in this section.

1.5.1 Outer Ear

The outer ear acts as a collector of sound. It does so mostly by scattering energy into the external ear canal, since its dimensions are not large enough, relative to the wavelengths of audible sounds, to permit efficient reflection. Since scattering increases with increasing frequency, the external ear is a more efficient collector for high-frequency sounds than for low-frequency sounds, and, because of its shape and position, is a more efficient collector for high-frequency sounds directed toward the face than it is for those directed toward the back of the head.

This fact can be demonstrated readily with the aid of a high-quality watch. The tick of such a watch is rich in high-frequency sound. If the watch is held beside the head, it will sound louder when in front than it does when behind the [illegible] This effect can be augmented by holding the cupped hand, concave forward, near the ear. In this case there is a considerable increase in loudness, when the watch is held in front, at the instant that the edge of the cupped hand is brought into contact with the ear. This is a resonance effect and is similar in origin to the roaring sounds heard when a seashell is held to the ear.

In other words, resonances in the external ear can amplify the loudness of some sounds or render them audible when they would otherwise be too faint to hear. Thus, measurements have been made[11] at a fixed position in space, later to be occupied by the head of the subject, of the sound pressure produced by a standard source. These measurements were repeated within the auditory canal when the subject's head was brought to the test position. The sound pickup was a narrow flexible tube whose free end could be inserted into the ear canal. The other end of this probe tube was used to drive a condenser microphone. For tones between 1 and 5 kilocycles, the pressure within the ear canal exceeded the free-field pressure. The greatest amplification, for tones with frequencies of 2 to 3 kilocycles, was between 9 and 15 decibels.

These measurements have recently been repeated in an echo-free chamber,[12] and differences between the free-field and auditory canal pressures as large as 20 decibels are reported for the case of 3-kilocycle tones. The resonance frequency and degree of amplification are found to vary from subject to subject and probably depend upon the shape and the size of the external ear. In other words, the ear canal behaves as though it were a closed tube with rigid walls and a length of about 2 centimeters. For comparison, the average depth of the ear canal is about 2.7 centimeters, and its diameter about 0.7 centimeter. Furthermore, it has been found[13] that the resonance frequency of the ear canal may be changed by filling it with hydrogen or coal gas. These observations lend support to the view[11] that reduction of deafness following surgery is often an incidental by-product and is largely due to a change in the resonance frequency of the outer ear, thereby improving response in some frequency regions and diminishing it in others.

1.5.2 Middle Ear

The ear drum is shaped somewhat like a loudspeaker cone and has very little inertia. It normally seals off the middle ear, which is essentially an air-filled cavity, from the external atmosphere. Differences of static pressure across the drum (such as are produced by rapid changes of altitude) constrain its motion and affect the responsiveness of the ear to sound.

The middle ear communicates with the throat through the *Eustachian tube* (see Figures 1 and 2). This tube opens during the act of swallowing, or yawning and thereby equalizes the pressures acting on the drum membrane. For a discussion of injuries produced by failure to maintain equality of middle ear and external pressures, in the case of submariners making practice escapes from a depth of 100 feet below the surface, see reference 15.

If the pressures exerted upon the round and oval windows were equal, they would cancel each other's effects, and motion of the basilar membrane would not occur. These pressures are unequal, however, because the force acting on the drum is transmitted to the oval window through the ossicles and to the round window by an air path. Since the drum has an area about twenty-five times that of the oval window, and since the ossicles are equivalent to a lever with a mechanical advantage of about 1.5, the pressure at the oval window is approximately forty times greater than at the drum. Similarly, the pressure at the drum is probably about ten times higher, for a 2-kilocycle tone, than that in the incident sound wave, owing to resonance amplification in the ear canal.

Since the motional impedance of air is less than that of the inner ear fluid, the middle and outer ears act as impedance matching devices. Measurements of the overall impedance of the ear show that the match to that of air is good in the region between 500 and 1,500 cycles per second.

Several mechanisms in the ear reduce somewhat the amplitude of response to loud sounds. In the first place, the ossicles cease to move as a unit when high-intensity sounds reach the ear; and for large displacements of the ear drum there is considerable sliding and friction at the joints in the chain of ossicles. Nonetheless, the amplitude of motion of the ossicles continues to increase with intensity.

In the second place, the ossicles are controlled, as shown by the arrows in Figure 2, by the action of two antagonistic muscles. These are tensed by reflex action in the presence of loud sounds and thereby diminish the amplitude of motion. Measurement shows that the protective action of the middle ear muscles is confined to frequencies below 1 kilocycle.

The middle ear mechanism becomes progressively less responsive to increasingly intense sounds. Indeed, the ossicles may disarticulate in the vicinity of the pain threshold. The tilts and relative positions of the ossicles tend to change as the amplitude of their motion increases. Thus, the mechanism of the middle ear is nonlinear, as illustrated in Figure 9. Here, the heavy curve represents the transfer characteristic of the middle ear. It will be seen that the transmitted displacements do not keep pace with the applied displacements.

For the sinusoidal input wave sketched on the vertical axis, the transmitted disturbance has the form shown on the horizontal axis. The output wave is seen to be flattened, with respect to the input, but the positive and negative half cycles are symmetrical about the horizontal line. This distortion of wave form, due to nonlinearity of the transfer characteristic, introduces higher frequencies into the transmitted disturbance which are odd harmonics (odd integral multiples) of the input frequency and which are not present in the input.

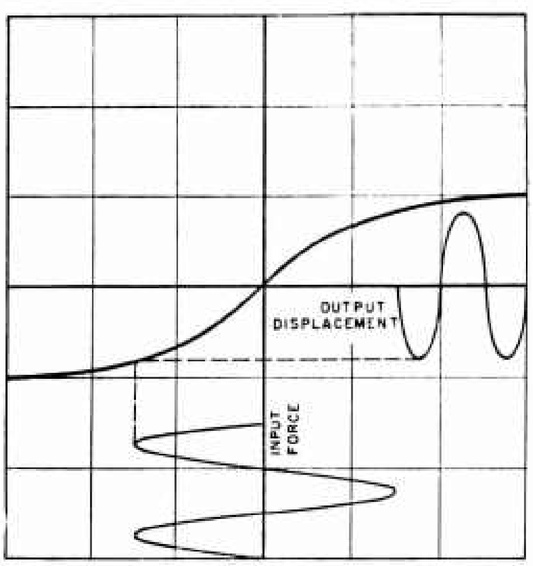


Figure 9. Nonlinear transfer characteristic.

When the transfer characteristic is asymmetrical, as well as nonlinear (see Figure 10), the transmitted wave contains the even harmonics of the fundamental input frequency in addition to its odd harmonics. Observations cited under "Aural Harmonics" in Section 2.1.2, indicate that the ear generates both the odd and even harmonics of single-frequency sounds incident upon it.

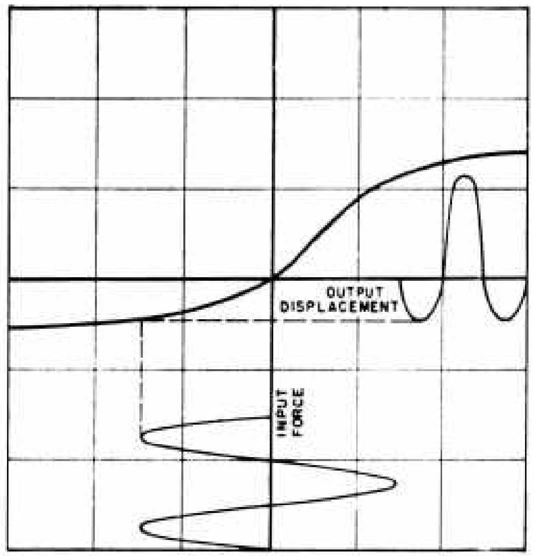


Figure 10. Asymmetrical and nonlinear transfer characteristic.

Although the ligaments from which the ossicles are suspended in the middle ear cavity generally prevent lateral vibrations of this chain of bones for low-intensity sounds, this is probably no longer true for large displacements and introduces a further source of distortion. It should be pointed out, in conclusion, that the middle ear is probably not the only locus of distortion. Thus, the ear drum very likely vibrates in segments, when the incident intensity is high, instead of moving as a unit, and the vibration frequencies of such segments will be higher than that of the drum membrane as a whole. Similarly, the tectorial membrane, even though it is relatively flaccid and does not hinder transmission of pressure changes to the basilar membrane, tends to load the latter differently during positive and negative half cycles of the incident disturbance.

Chapter 2

BEHAVIOR OF THE EAR

THE FUNDAMENTAL BEHAVIOR of the ear depends on the structural factors discussed in the preceding chapter and sheds further light on the problems already discussed. The present analysis is more immediately concerned with what rather than how we hear and gathers together data which will be useful in connection with problems studied later in this volume. The chief points around which the present chapter is developed are: the faintest wanted sounds audible in the presence and absence of interfering, or unwanted, sounds; sensations of pitch and loudness; the auditory effects of sounds containing many component frequencies; and the sensed effects produced by varying the durations of the sounds and their components.

2.1 THRESHOLD FOR TONES

Four major types of threshold may be distinguished. Of these, three kinds of absolute threshold are described in Section 2.1.1. (An absolute threshold is the critical level of a sound when no background is present.) A fourth variety, the threshold in the presence of masking tones, is defined and discussed in Section 2.1.2. Thresholds for pure tones are considered first in the present section, whereas complex sounds are considered in Section 2.3.

2.1.1 Absolute Threshold

The first of the three absolute thresholds is called the audibility threshold, or the absolute audibility threshold, and refers to the smallest rms pressure which a pure tone must have in order to be heard by an average, normal ear, provided the surroundings are quiet, so that the listener experiences no interference from unwanted sounds. The second kind of threshold refers to excitation of the hearing mechanism by methods other than varying the air pressure in the ear canal. The third kind of threshold, variously termed the threshold of pain or the threshold of feeling, refers to sensations of a nonauditory character which are produced by very intense airborne sounds.

It should be understood that all these thresholds actually refer to the midregions of zones within which certain phenomena become more or less well defined and do not specify anything more than average values. Consequently it is important to state in all cases the conditions of measurement and the variability of observer response.

AUDIBILITY THRESHOLD

Figure 1 shows the audibility threshold at various frequencies for three experimental situations. These thresholds indicate the response of acute ears when the static pressure is about 760 millimeters. Since these results were obtained from a large number of observers, individual peculiarities, which are characteristic of outer ear resonances, have been averaged out. It is interesting to note that the threshold of hearing is much the same at static pressures corresponding to an altitude of 36,000 feet above sea level.[1]

The general shape of the threshold curve implies that the overall response of the ear mechanism as a whole, consisting of the outer, middle, and inner portions coupled together, resembles that of a broadly tuned resonator whose peak response lies in the interval between 1 and 7 kilocycles. That this response characteristic is not associated exclusively with only one portion of the ear is indicated by the effect of surgical changes in the shape of the outer ear canal (see Section 1.5.1) and also by the following set of observations made in cases in which the experimenter had direct access to the middle ear cavity.[2] Increasing the effective stiffness of the chain of ossicles by inserting a tuft of compressed cotton between them increased the acuity of the ear to high frequencies. This would be expected from the fact that the natural frequency of a mechanical resonator increases when the stiffness, or coefficient of restoring force, is increased (see equation (1) in Chapter 1, for example). In addition, when the subjects reclined on their sides and a drop of mercury was placed on the round sponse is mass controlled at high frequencies (response falls for high accelerations due to inertia) and stiffness controlled at low frequencies (when inertia is unimportant, the role of stiffness is more prominent). For frequen-

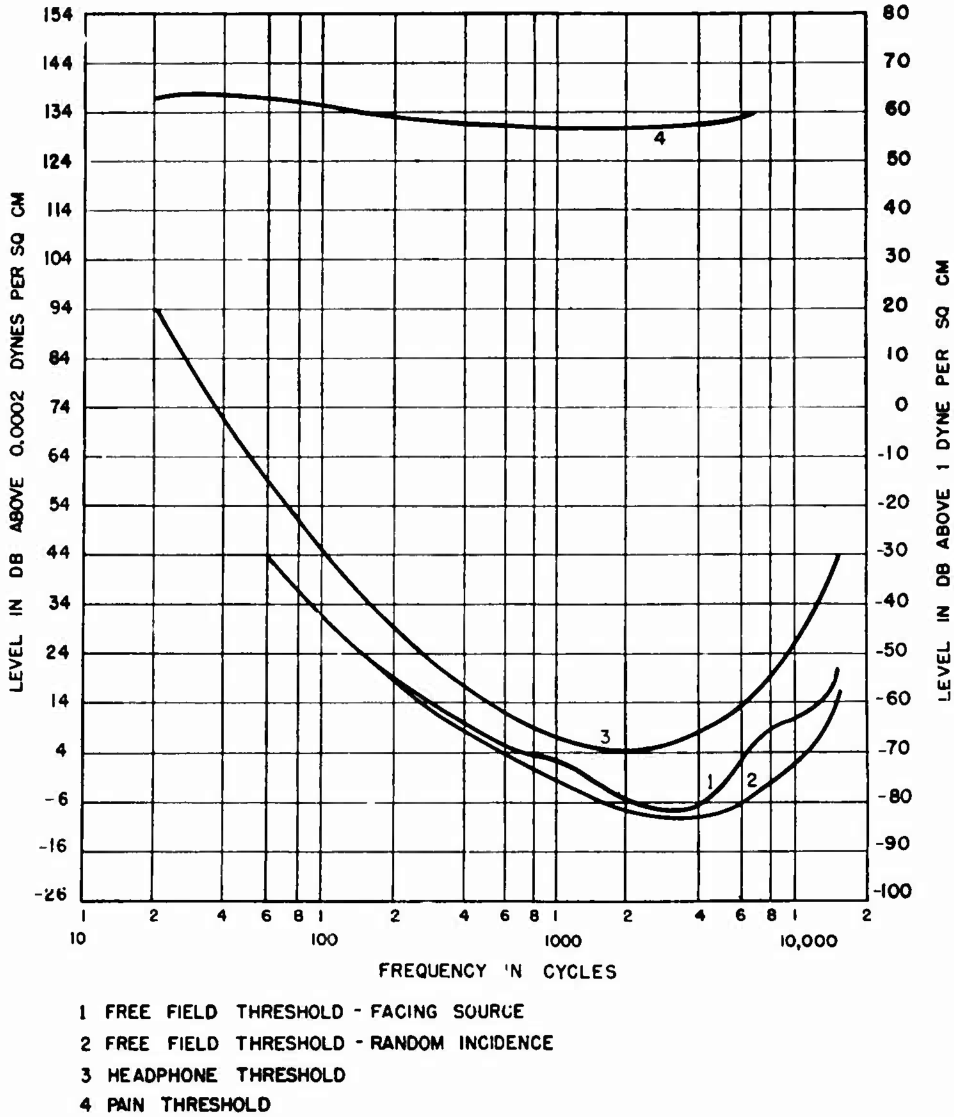


FIGURE 1. Audibility thresholds at different frequencies. (Courtesy John Wiley and Sons.)

window, the increase of mass increased acuity for low frequencies. Similar results have been obtained in animal experiments in which tissue was grafted over the round window and the cochlear potentials measured before and after the graft was made.[3]

The results of impedance measurements made upon the ear[4] also indicate that its response is mass controlled at high frequencies

cies at which the effects of stiffness and mass are balanced, or nearly so, the system is in resonance.

Curve 1 in Figure 1 was obtained by determining the pressure at a stated position with respect to a source of single-frequency sound. After the level of the sound in this free field had been measured, the heads of the observers were placed at the reference position, facing the source, and their ability to hear the generated tones was noted. This curve, as well as curves 2 and 3 in the same figure, shows that the faintest audible tones are those with frequencies in the neighborhood of 3 kilocycles. At this frequency, the faintest audible sounds have a free-field pressure of about 80 decibels below 1 dyne per square centimeter, as shown by the right-hand scale. This corresponds to an rms pressure p_0 of 10^{-4} dyne per square centimeter, in other words, $20 \log (1/10^{-4}) = 80$ decibels. Since the intensity is equal to $p_0^2/\rho c$, where ρ, the density of air at 760 millimeters and 20 C, is 1.2×10^{-3} gram per cubic centimeter, and c, the velocity of sound propagation under the same conditions, is 3.4×10^4 centimeters per second, it follows that the power supplied to the minimum audible field amounts to $(10^{-4})^2 / [(1.2 \times 10^{-3}) (3.4 \times 10^4)]$, or 2.5×10^{-10} erg per second per square centimeter. This is equal to 2.5×10^{-17} watt per square centimeter.

Since the intensity (2.5×10^{-10} erg per second per square centimeter) may also be expressed as the quantity $2\pi^2 f^2 A^2 \rho c$, where f, the frequency of the tone, is 3 kilocycles, A is the peak displacement amplitude, and the other symbols have the significance already stated, the maximum displacement of an air particle in the threshold sound field may be readily computed. Thus

$$2.5 \times 10^{-10} = 2\pi^2 (3 \times 10^3)^2 A^2 (1.2 \times 10^{-3})(3.4 \times 10^4)$$

or

$$A = 2 \times 10^{-10} \text{ cm.}$$

Attempts to measure the displacement of the eardrum at threshold have given roughly this same value at frequencies between 1 and 10 kilocycles. Thus the maximum displacement at threshold is at most of molecular dimensions.

It may also be shown[5] that the intensity of the noise generated by the random thermal motion of the air particles is of the order of the intensity of the faintest audible 3-kilocycle tone. Thus, a further increase in the acuity of the ear in the region between 1 and 5 kilocycles would not make it possible to hear fainter wanted sounds, since these would be drowned out by the thermal noise in the air and in the structures of the ear.

The irregular shape of curve 1 in the regions centered at 1 and 6 kilocycles is due to the fact that the intrusion of the observer changes the distribution of energy in the sound field. These irregularities in curve 1 are associated with the effects of diffraction around the head and shoulders of the observer. The diffraction pattern changes with the frequency of the source and the orientation of the head relative to the source. Thus, when the sound is propagated toward the side of the head (that is, directed upon the ear) instead of toward the face, the minimum audible field for a 3-kilocycle tone is nearly 100 decibels below 1 dyne per square centimeter and is proportionately lower than indicated by curve 1 for frequencies above and below 3 kilocycles.

Curve 2 shows the effect of correcting for diffraction and has been computed by assuming that the head is immersed in a random, or nondirectional, sound field. Since the sound shadow cast by the head is negligible for low frequencies, curves 1 and 2 do not diverge appreciably before the source frequency exceeds 500 cycles per second. In general, curve 2 is lower than curve 1 because more sound reaches the ear at an optimal orientation in the case of the random sound field.

Curve 3 indicates the absolute threshold determined with a headphone applied to one ear. This monaural threshold was obtained by determining the pressure in the enclosure formed by the tightly applied headphone. Such pressure measurements are very difficult to make at the low values required. The threshold pressure was actually found by a theoretical computation, based on knowledge of the volume of enclosed air and the displacements of the headphone diaphragm. This method involves two assumptions. In the first place, leakage of air along the headphone seal is neglected, which is probably not justified at the lowest frequencies. In the second place, the method assumes that the sound pressure is a linear function of the electrical current fed into the headphone, since the displacement of the headphone diaphragm is assumed proportional to the current; this assumption may not be legitimate for the entire range of sound pressures used.

The shape of curve 3 is in general the same as that of curve 2, but the monaural pressure measurements are, on the average, about 15 decibels above the binaural field measurements. There are many factors which tend to produce this shift, but it is not known whether they account for all the difference between curves 2 and 3. Thus, scattering of sound into the ear canal by the outer ear structures (see Section 1.5.1) increases with increasing frequency, and resonances in the outer ear also tend to amplify the pressure at the eardrum with respect to the pressure measured in the free field. Furthermore, unless echoes from the walls of the test enclosure are completely eliminated, they may return toward the sides of the head in a manner different from that by which they reach the pickup used in the free-field calibration. Such echo effects have been found capable of depressing the free-field threshold by 3 to 5 decibels.

Differences between curves 2 and 3 at frequencies below 1 kilocycle are not likely to be due to the factors mentioned above. On the other hand, there are at least three effects known to occur at these low frequencies which would be expected to produce differences between the headphone and free-field thresholds; namely, reflex tensing of the middle ear muscles, physiological noise, and acoustic leakage. Reflex tensing of the middle ear muscles occurs in the case of loud sounds, and this tension increases the acoustic impedance of the ear for frequencies below 1 kilocycle. Since the headphone pressure measurements are made by determining the electrical energy fed to the receiver at threshold and also about 60 decibels above threshold, the possibility that the middle ear muscles are in different states of tension at these two levels means that the acoustic power used to move the eardrum is not a linear function of the electrical power fed to the headphone.

The term "physiological noise" refers to mechanical vibrations set up in the space enclosed by the tightly fitting headphone due to pulse actions, motion of air through the head cavities during the act of breathing, and similar phenomena. Such noise would tend to drown out the wanted sound and elevate the low-frequency threshold.

Even when the headphone cap fits fairly tightly, leakage of air along the seal is likely to be greater during the long period of a low-frequency tone than during the short period of a high-frequency sound. This leakage will tend to make the actual pressure change in the ear canal smaller than the computed changes.

Two additional aspects of these threshold measurements should be mentioned at this point: (1) the difference between monaural and binaural thresholds and (2) the element of observer variability. It has been shown[6] that the total energy required to reach threshold by means of a tone led to one ear is equal to the sum of the energies of tones led independently to both ears. Furthermore, this equality is independent of the numerical ratio in which the energy is divided between the two ears. In fact, the frequencies as well as the intensities of the tones led to the two ears may differ from each other over a fairly wide range and the binaural threshold will still be reached at the same value of the total intensity, provided the two ears are fairly well matched with respect to acuity. These observations imply that the associative centers in the brain integrate the nerve discharges produced by stimulation of both ears, and there is other evidence supporting this view (see Section 8.1.3). In addition, the results just quoted indicate that the binaural threshold should lie some 2 to 3 decibels below the monaural threshold.

As to observer variability, it should be noted that the position of the threshold varies from day to day for a given observer, and even from minute to minute. Thus, if a sustained tone is presented to an observer at a level near threshold and he is told to press a key whenever the tone is audible and to release it when the tone is inaudible, he will in general press the key intermittently. Similarly, when groups of tonal pulses with intensities near threshold are presented and the observer compares the duration and succession of his acoustic sensations with the actual duration and grouping of the tonal stimuli, it is found that he is, in general, unable to hear all the pulses in a group, and that the sensed duration of audible pulses is less than their actual duration.[7] The same effect was observed when the threshold had been artificially elevated by strong stimulation. It follows that the level of threshold must be defined in statistical terms. Unless experimental measurements of threshold refer to the same probability of perceiving the test sound, such measurements are not strictly comparable, and in extreme cases may vary from each other by as much as 10 to 15 decibels.

Because of the spread in the measurements, an average between curves 2 and 3 is generally adopted, and the threshold for a 1-kilocycle tone is usually defined as 74 decibels below 1 dyne per square centimeter, in other words, 2×10^{-4} dyne per square centimeter. In many types of work this rms pressure value is used as the reference standard, and therefore corresponds to a level of 0 decibels (see left-hand ordinate in Figure 1). The convenience of this standard is largely associated with the fact that most sound levels of interest will have positive values.

The decibel scale expresses the value of a ratio. In underwater acoustics, the conventional standard which is used to form this ratio is an rms pressure of 1 dyne per square centimeter; in atmospheric acoustics, it is customary to use 2×10^{-4} dyne per square centimeter as the reference standard. Since the present discussion is concerned with hearing as well as with underwater sounds, both scales are given wherever they are relevant. The air standard is used in all scales on the left of the diagrams; the water standard is printed at the right. Any given pressure p is 74 decibels higher on the "air" scale than it is on the "water" scale.

Bone Conduction Threshold

In some diseases of the ear, surgical removal of the ossicles (except for the fragment which seals the oval window) is necessary. While loss of the ossicles raises the audibility threshold by about 60 decibels, it does not produce complete deafness. The available evidence indicates that the inner ear mechanism is stimulated in such cases by vibrations of the skull structure in response to the incident airborne sound. Thus, as long as the inner ear and the auditory nerve remain intact, and the osseous pathway has not been impaired by skull fracture, it is always possible to deliver sound vibrations at sufficient intensity to permit useful hearing.

Bone conduction thresholds are usually determined by applying the end of a vibrating rod to the chin, the forehead, or the mastoid process. For individuals with normal hearing, the bone conduction threshold, plotted as displacement amplitude versus frequency, continues to be concave upward as in Figure 1; but the curve tends to be nearly flat between 1 and 6 kilocycles,[8] owing to the fact that there is virtually no resonance amplification such as is normally introduced by the outer ear. At threshold, for a frequency of 800 cycles per second, the amplitude of motion of the forehead immediately adjacent to the vibrating rod has been found to be 3.5×10^{-9} centimeter in the case of normal ears[9] and is probably nearly the same when the ossicles have been removed. The necessary displacement amplitude falls to 5×10^{-10} centimeter, if the ear canal is closed by inserting the finger into it. Under these circumstances, the sealed-off air is set into vibration,[10] the eardrum moves, and energy is transmitted to the cochlea through the ossicles.

Motion of the basilar membrane occurs in the case of bone-conducted vibration because of two asymmetries in the structure of the cochlea. Thus, compression of the semicircular canals increases pressure to a greater extent in the upper gallery of the cochlea than in the lower (see Figures 2 and 6 in Chapter 1). Similarly, the round window yields more readily than the oval window when the cochlea is compressed, and this also tends to establish a pressure gradient across the basilar membrane, directed from the upper to the lower gallery.

The fact that bone conduction occurs means that airborne sounds produced in noisy locations will interfere with the audibility of signals brought to the ear by means of headphones. Under ideal conditions, therefore, the insulation afforded by headphones against unwanted airborne sounds is not likely to exceed 60 decibels and will usually be less.

Since air and bone conduction of a given tone stimulate the same portion of the basilar membrane, it is possible, by adjusting the phases and intensities of simultaneous osseous and atmospheric stimuli, to make the subjective sensation disappear.[11]

PAIN THRESHOLD

Tests conducted with airborne sounds of frequencies between 1 cycle and 8 kilocycles show that nonauditory sensations — tactile sensations, feelings of pressure, thrusting, tickling, itching, and burning — appear whenever the intensity of the stimulus tone reaches a value of about 130 decibels above 2×10^{-4} dyne per square centimeter, as indicated by curve 4 in Figure 1. Since permanent injury results from continued exposure to sounds of higher intensity, the useful range of auditory stimuli lies in the area defined by the threshold of hearing and the threshold of pain.

2.1.2 Masked Threshold

It is a common experience that the level of the voice must be raised if the speaker wishes to be heard at a noisy location. This phenomenon is termed *masking*. Quantitative studies of masking have provided much useful information concerning the behavior of the ear; hence the masking effect of one tone upon another will be described in the present section.

The unwanted sound which produces the masking will be called the background (or the masking tone) and the wanted sound will be called the signal (or masked tone). Signal or background, or both, may be atonal sounds, and the masking produced by and upon such sounds is discussed in subsequent sections.

In general, the background must be audible, otherwise it will not produce masking. Since the effect of a masking tone is usually to raise the audibility threshold of the signal, it is customary to define the amount of masking by specifying the difference in decibels between the level of the signal when it is (1) presented to the ear at a just audible level without interference from masking, that is, the level at the absolute audibility threshold, and (2) presented to the ear at sufficient strength to become just audible in the presence of the masking tone, that is, the level of the masked threshold.

The level of the audibility threshold depends only on the acuity of the ear; the level of the masked threshold depends also upon the properties of the background. In effect, the masking sound produces a temporary deafness to the signal by shifting its audibility threshold.

A group of charts indicating the threshold shifts produced by masking are shown in Figure 2. The frequency of the masking tone is indicated at the top of each chart, and a number specifying the sensation level, or relative intensity, of the masking tone is used to label each masking curve. By *sensation level* is meant the number of decibels by which the level of the masking tone exceeds its own threshold level. From Figure 1, it will be seen that when a 200-cycle tone has a sensation level of 80 decibels, it is about 110 decibels above 2×10^{-4} dyne per square centimeter whereas a 1,000-cycle tone at a sensation level of 80 decibels is only 80 decibels above 2×10^{-4} dyne per square centimeter. Thus, the sensation level of a tone does not directly state either its objective intensity or, as will be shown later, its subjective loudness; nonetheless, it is a precise and useful concept.

The frequency of the masked tone is shown on the horizontal scales in Figure 2, and the threshold shift of the masked tone (the sensation level of the signal just audible in the presence of the indicated background) is given by the vertical scales. The threshold shifts were obtained by testing monaurally the ears of groups of normal observers.

The signal and background tones were generated by electrical oscillators, freed of their harmonic content by filtering, and fed to headphones. The power input required to reach the audibility threshold was determined individually for each of the signal and background tones. In addition, a mixture of a background tone (at a fixed intensity) and one signal (whose intensity could be varied gradually) was applied to the headphones. The threshold shift was then obtained by noting the power supplied by the signal oscillator when the masked signal could be heard half the time by the various observers, or by the same observer in half the tests.

This averaging procedure is necessitated by the fact that the masked threshold is not infinitely sharp. As the signal-to-background ratio is gradually increased the signal undergoes a transition (extending over a range of about 12

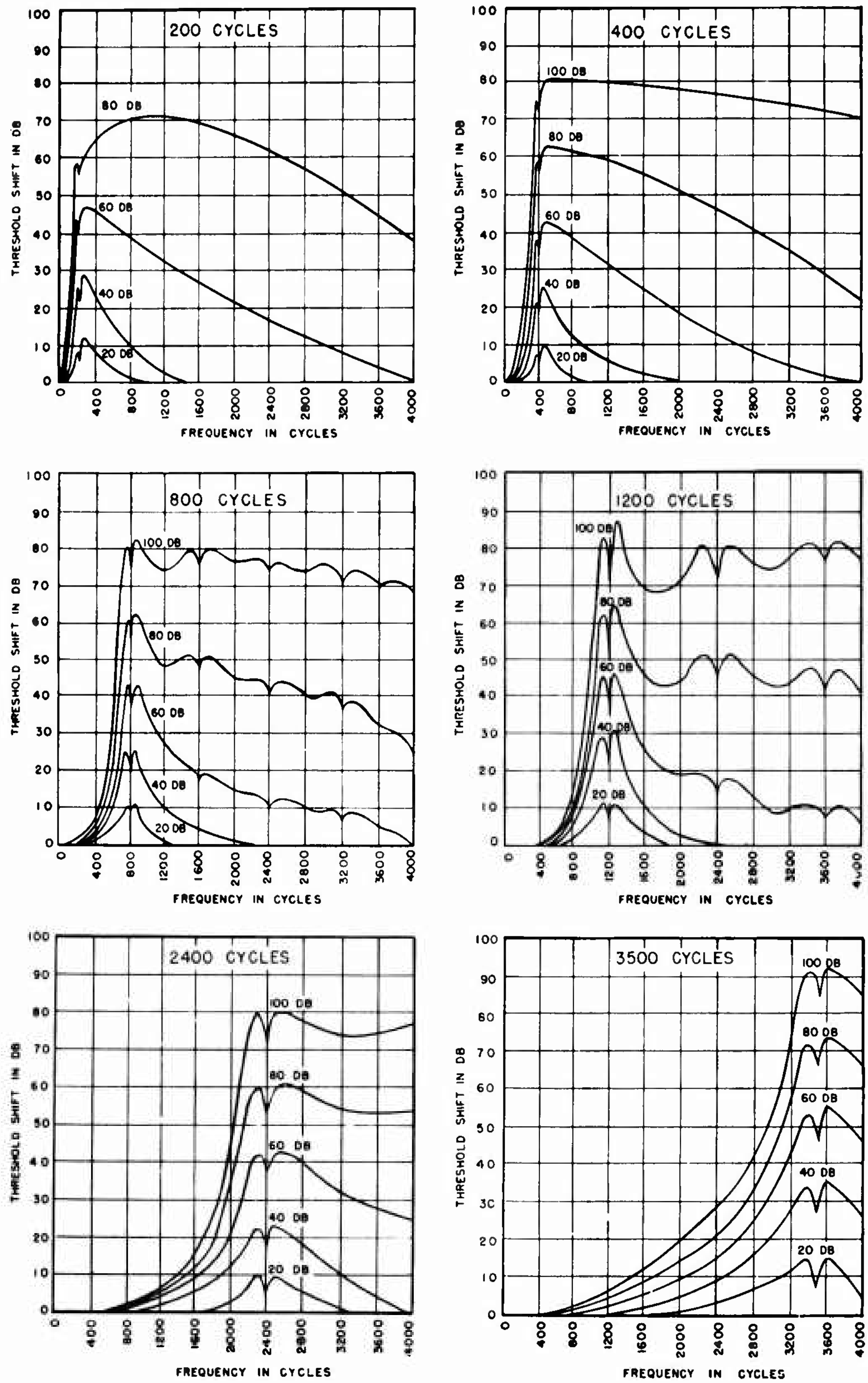


FIGURE 2. Masking of tones by tones (200, 400, 800, 1,200, 2,400, and 3,500 cycles). (Courtesy D. Van Nostrand Co.)

decibels; see Section 4.1.4, for example) from "inaudible," through "audible some fraction of the time," to "audible all the time." It is therefore convenient to adopt the signal-to-background ratio at which the signal is audible in one-half, or some other stated fraction of the trials as defining the masked threshold. When defined in this manner, the level of the masked threshold for a specific signal-background combination can usually be duplicated to within 1 or 2 decibels in successive determinations.

Three general conclusions may be drawn from these charts. First, the threshold shift is greatest at frequencies adjacent to that of the background. Secondly, the threshold shift at remote frequencies increases when the sensation level of the background is raised, especially when it is larger than 40 decibels. Finally, the threshold shifts at remote frequencies are more severe in the direction of high frequencies.

It should be stated at this point that the masking curves obtained are significantly different from those shown in Figure 2 when the masking tone is introduced into one ear and the masked tone into the other. In general the level of the masking tone presented to one ear must be about 50 decibels higher than that shown in Figure 2 in order for it to produce the indicated threshold shifts of the masked tone led to the other ear. The available evidence indicates that this binaural masking effect does not occur in the brain, but rather that bone conduction occurs, and the masking produced in the ear to which the signal is presented arises from the bone-conducted sound transmitted across the head from the opposite ear. The attenuation factor of 50 decibels agrees with the bone-conduction data described in Section 2.1.1. This conclusion agrees with observations made on individuals with one normal ear and one totally deaf ear. When a telephone receiver is held to the deaf ear in such cases, the level of the presented tone must be raised approximately 50 decibels above normal threshold in order to be heard. That this tone is actually heard in the normal ear is indicated by the fact that its loudness is increased by inserting a finger in the canal of the normal ear.

It should also be added that the masking data shown in Figure 2 apply only when signal and background are presented simultaneously and have a duration of between 1 and 10 seconds. For presentation intervals of different length and for successive rather than simultaneous presentation of the two sounds, other factors strongly influence the results.

For example, it has been observed[12] that the threshold shift of a 2,250-cycle signal, masked by a 1-kilocycle background is not independent of time. In these experiments, the 1-kilocycle background was presented continuously and the masked threshold of the signal was determined at some known time after the background was turned on. When the signal threshold was determined 10 seconds after the onset of background, the shift amounted to about 12 decibels. The threshold shift grew progressively smaller with time and remained substantially constant — at a value of 4 decibels — for all determinations made more than two minutes after the background was first turned on. During the entire test period the intensity of the 1-kilocycle background was maintained constant, at a sensation level of about 50 decibels.

That this diminution of threshold shift was not due to practice in making successive determinations is indicated by the fact that the same low value of 4 decibels was found even when the first test was made after the observer had listened passively to the uninterrupted background for about five minutes. Since the sensed character of the sustained background changed with time — it is described as growing "more mellow" and "less insistent" — the experimenters concluded that the reduction of masking was due to fatigue. According to this view, the acoustic nerve became less responsive to the sustained background and hence relatively more responsive to the signal.

It seems to follow that (1) the signal-to-background *ratio* corresponding to the masked threshold would be independent of time if the signal were also a sustained sound, and (2) the characteristics of masking are determined by the nature and behavior of the auditory nerve tract as well as by the dynamics of the cochlea.

Adjacent and Remote Masking

If the basilar membrane had perfect selectivity — if, in other words, the motion of a particular auditory string could be excited only when a tone of its own frequency reached the ear — the masking effect of any background tone would be confined to signals of identical frequency. As Figure 2 shows, the tuning is not perfect. Most of the masking does seem to be confined to signals with frequencies immediately adjacent to that of the background, especially for backgrounds of low intensity, and this explains why it is often helpful to change the pitch of the voice when speaking in noisy surroundings. This type of masking is called *adjacent masking*. However, as the sensation level of the background increases, its masking effect extends to signals with frequencies more and more remote from that of the masking tone. In other words, when a tone has a sufficiently high intensity it produces motion of a large fraction of the basilar membrane. This masking of distant frequencies is called *remote masking*. These two concepts, remote and adjacent masking, are helpful in discussions of masking data.

Beats

The reason that the masking curves show a dip, or relative minimum, instead of having a single peak at one frequency of the background, where this sound produces its maximum stimulation, is that beats are heard when the signal and background tones differ in frequency by a small number of cycles per second. The audibility of the beats makes it much easier to detect the presence of the signal, since they produce easily recognized changes of loudness. When the signal is very brief, the beats cannot be heard, and the masking curve has a single maximum at the frequency of the background (see Figure 5 in Chapter 10).

The occurrence of audible beats when two tones of comparable intensity are sounded together indicates that each tone stimulates a finite segment of the basilar membrane. When these segments of the membrane overlap, the net disturbance due to the simultaneous action of both tones will vary with time, being greatest when the two stimulated patches move in phase and least when they are of opposite phase. Thus, the net disturbance of the membrane depends significantly on the relative phases of the two tones, but only when these tones are sufficiently close in frequency so that they can stimulate the same portion of the basilar membrane. In other words, the occurrence of beats indicates that the ear is not a perfect analyzer; that is, it does not have infinitely sharp tuning.

This fact also explains the observation that the threshold shift is sometimes negative, that is, the signal may become audible below its absolute threshold when the presence of a background tone, sufficiently close to it in frequency, causes the total stimulation of the patch Δs to reach threshold.

According to this view, the masking curves would bear a close resemblance to simple resonance curves, at least for background levels below 50 decibels, if the loudness changes due to beats did not produce a dip at frequencies adjacent to that of the masking tone. The reason for changes in the appearance of the masking curves when the sensation level of background exceeds 50 decibels is the fact that the nonlinearity of the ear introduces harmonics of the masking tone, as discussed on page 24 ff.

Thus, Figure 3 has been redrawn from Figure 2 in order to show the type of curve which would be obtained if no assistance were afforded by beats. The peak level of the hatched portion of each of these curves at the frequency of background has been set equal to the sensation level of the background. This procedure assumes that ability to hear the signal (rather than beats associated with its presence) is possible if the stimulation of the basilar membrane produced by the signal is equal, for at least one patch of the membrane, to the stimulation produced there by the background. The assumed identity of masking and stimulation has proved to be a useful principle in the solution of a variety of problems. The peak levels of darkened areas corresponding to harmonics of the masking tone have been derived from measurements described on page 24 ff.

It will be seen that the height and base width of each of the hatched areas increases with the intensity of the masking tone. Both of these circumstances are associated with the audibility of beats. Consider, for example, the height of the hatched area (24 decibels from the bottom of the dip to the top of the peak) for the case of an 800-cycle tone at a sensation level of 60 decibels. If the enhanced audibility of signals which differ from background by 2 to 3 cycles per second is due to the perception of loudness changes, the improvement of 24 decibels should be determined by the fact that under these circumstances (see Section 2.2.2) the smallest change of intensity which the ear can detect amounts to about 25 per cent. In other words, when the intensity of an 800-cycle tone at a level 60 decibels above threshold is periodically varied three times per second so that its highest intensity is 1.25 times its lowest intensity, the fluctuation of level is just detectable.

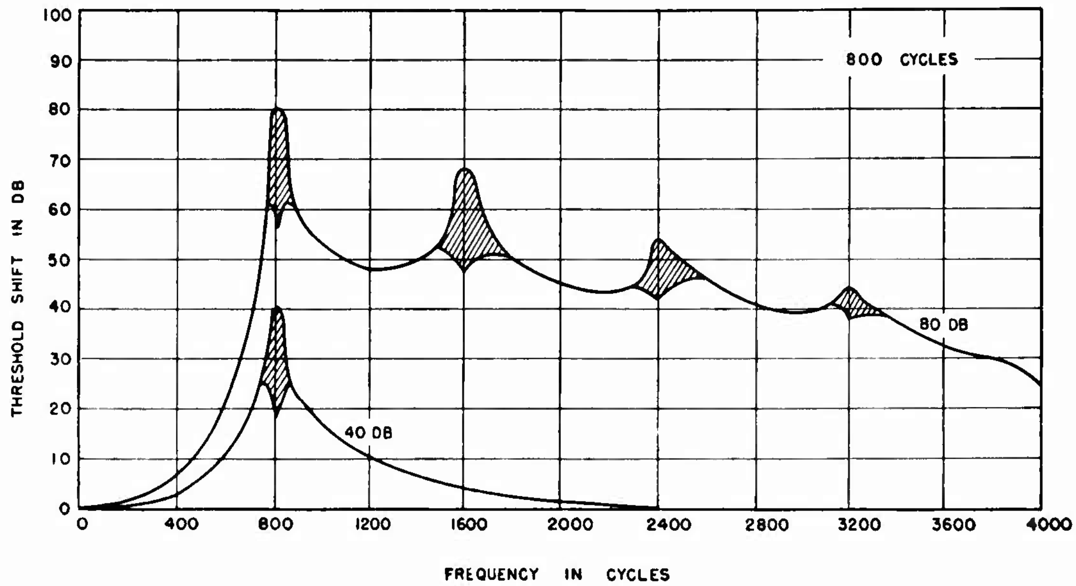


FIGURE 3A. Assumed stimulation of the basilar membrane at different sensation levels for an 800-cycle tone.

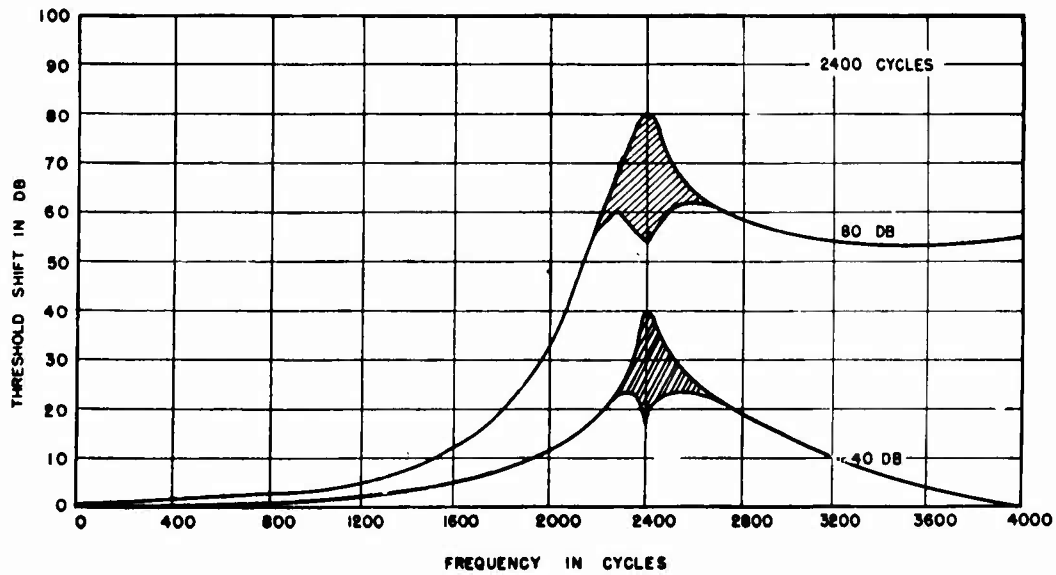
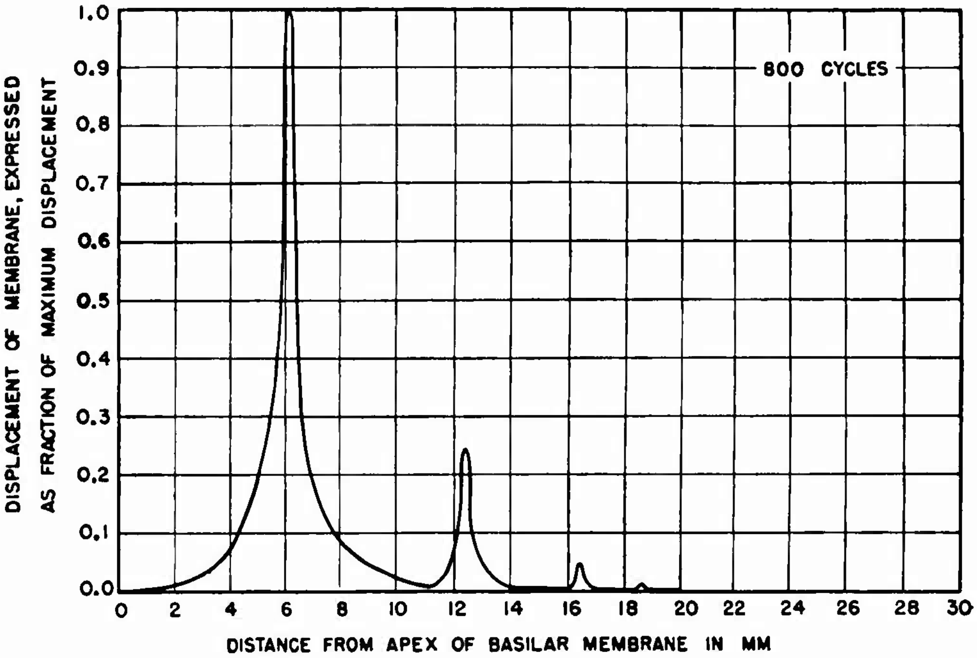
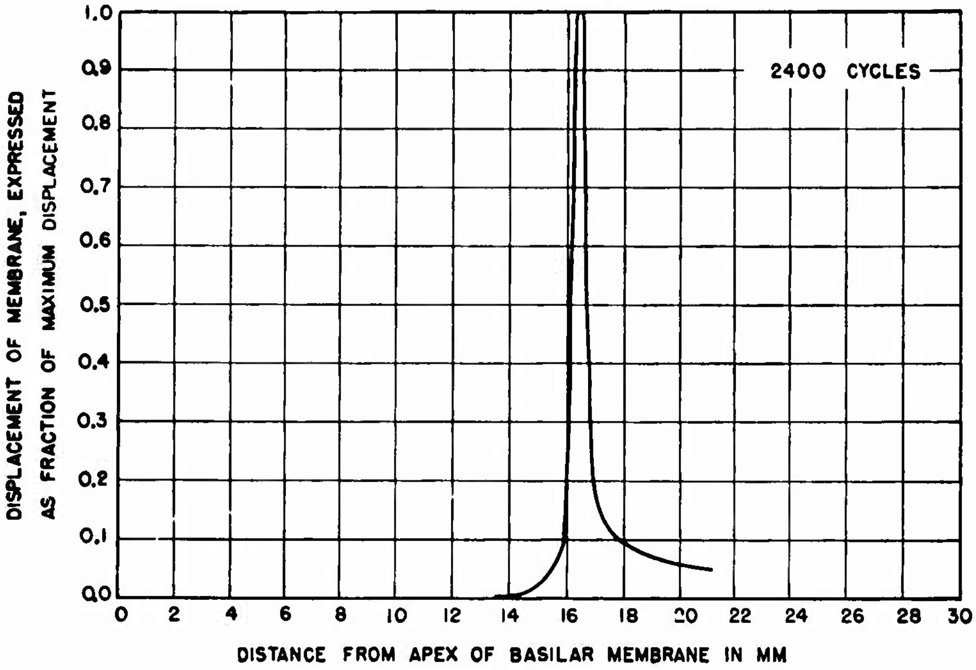


FIGURE 3B. Assumed stimulation of the basilar membrane at different sensation levels for a 2,400-cycle tone.

This fluctuation may also be considered to arise from the mixture of an 800-cycle back-



A



B

FIGURE 4. Displacement patterns produced by an 800- and a 2,400-cycle tone, each at a sensation level of 80 decibels. These patterns have been computed from Figure 3.

ground with a 797-cycle or an 803-cycle signal. Let the peak pressure of the 800-cycle background be P and the peak pressure of the signal be p. Then the peak intensity of the mixture will be proportional to $(P + p)^2$, and the minimum intensity to $(P - p)^2$. When the ratio

$$\frac{(P + p)^2}{(P - p)^2} = 1.25,$$

the periodic fluctuation of intensity will be just audible. Therefore, $p/P = 0.052$, and 20 log $0.052 = -25.6$ decibels, which is in good agreement with the observed fact that the level of the just audible signal is 24 decibels below the level of background when the two tones have nearly identical frequencies.

The increase in the height of the hatched area with sensation level agrees with the observation (see Section 2.2.2) that the smallest audible change of intensity is a smaller fraction of the comparison intensity when the latter is raised. Similarly, the decrease in the height of the hatched area with increasing difference between the frequencies of signal and background (and hence, an increase in the rate of beating) is due to the fact that it is more difficult to detect intensity fluctuations when the rate of fluctuation increases (see Figure 16). In addition, the edge of a hatched area marks the point where the rate of beating is too great and the overlap of the patches of membrane stimulated by signal and background too small to permit detection of the signal by means of periodic changes of loudness.

It should be noted that the curves in Figure 2 are plotted to a logarithmic ordinate in order to reduce them to a manageable scale. Furthermore, the abscissa gives the frequency, rather than the position along the basilar membrane. If it is assumed that the maximum displacement produced by the background at any point on the membrane is proportional to the threshold shift (expressed in terms of pressure) and that the position on the basilar membrane corresponding to any frequency is given by equation (5) of Chapter 1, Figure 4 may be drawn. This figure shows the estimated disturbance patterns produced by an 800- and a 2,400-cycle tone, each at a sensation level of 80 decibels, at the moment when the centers of the vibrating patches reach their maximum displacements simultaneously. Since the data on which the 2,400-cycle figure is based do not extend beyond 4 kilocycles, no harmonics are shown in this case.

The width of a resonance peak is usually measured between points at which the response is half the maximum value. Using this criterion, it will be seen that the width of a peak at any frequency is about 6×10^{-2} centimeter. The width of the region stimulated by a harmonic at any frequency has about this same value. It may be noted that this value is probably an overestimate, since the widths of the peaks shown here are based on the presence of beats and thus involve the overlapping of two stimulated patches of membrane. Hence, the stimulation pattern produced by a single tone is probably more nearly 3×10^{-2} centimeter in width. As shown by Figure 3, the widths of the peaks increase somewhat with intensity. The indications that patch width is relatively independent of stimulus frequency and that it has a numerical value of about 3×10^{-2} centimeter, or 1 per cent of the total length of the basilar membrane, are in good agreement with the evidence discussed in Sections 2.2.1 and 2.3.

By way of conclusion, it is worth describing the sensations produced by two tones of frequency f_1 and f_2 that produce beats. When the frequency difference between the tones is less than 1 cycle per second, the sensation is that of a succession of smooth increases and decreases in loudness. This loudness variation is most noticeable when the rate of beating is between 2 and 3 cycles per second (see also Section 2.2.2) and the intensities of the stimulus tones are equal. At rates of between 2 and 7 beats per second the sensed pitch of the beating complex seems to lie midway between the two stimulus tones and is therefore called the intertone, i.e., the maximum stimulation of the membrane occurs at a point between the centers of the patches disturbed by each of the beating tones.

At about 7 beats per second, the periodic variation in loudness ceases to have a smooth growth and decline, the beats are heard as a series of intermittent impulses, and the pitches of the two stimulus tones are heard in addition to that of the intertone. As the rate is increased to about 15 per second, the intermittence associated with the individual beats begins to sound like a flutter. For rates exceeding 20 per second, the intertone ceases to be heard, leaving only the two stimulus tones, and the sensation is associated with a "roughness" or a "rattle" which makes such combinations extremely dissonant, especially in the region of 30 to 70 beats per second.

For sufficiently large frequency separations, the roughness ceases to be heard. The cessation of roughness is associated with the fact that the disturbed patches of the membrane cease to overlap to any significant extent. These patches have approximately the same widths for all stimulus frequencies, but the frequency interval contained in each patch increases with the frequency of the stimulus tone (see equation (5) of Chapter 1 or Figure 9). Hence, the frequency difference ($f_2 - f_1$) between the stimulus tones for which roughness vanishes increases as the frequencies f_1 and f_2 are increased, as shown by Table 1.

TABLE 1. Roughness due to beats.

Frequency of lower tone in cycles	Number of beats per second Greatest roughness	Vanishing roughness
96	16	41
256	23	58
575	43	107
1707	84	210
2808	106	265

The sensitivity of the ear to beats has received an interesting application in the detection of methane (or fire damp) in coal mines. Since methane is less dense than air, sound is transmitted through it with greater velocity; hence, an organ pipe filled with methane, or an air-methane mixture, has a higher pitch than a similar pipe filled with pure air. When two such pipes, one of them filled with the suspected mine air and the other with pure air, are sounded simultaneously, beats will be heard when significant quantities of methane are present. The concentration of the methane can be estimated from the rate of beating.

Beats may of course occur directly between the stimulus tones, or between such a tone and an aural harmonic of the second, or between the aural harmonics of both. The nature of aural harmonics is discussed in the following section.

AURAL HARMONICS

When a pure tone of high intensity stimulates the ear, the sensed sound contains frequencies which are harmonics of the stimulus tone and which are not present in the objective stimulus. Thus, when an intense 200-cycle tone is presented, the observer can also hear tones of 400 and 600 cycles per second.

The aural harmonics of the stimulus have been investigated by means of probe tones, which give audible beats when their own frequencies are adjacent to those of the aural harmonics. In fact, this is essentially the meaning of the dips in the masking curves at multiples of the background frequency.

The absence of dips at multiples of the masking frequency for the 200-cycle and 400-cycle background tones in Figure 2 should not be construed to mean that aural harmonics do not occur for low-frequency tones. Actually, aural harmonics are quite prominent for the low-frequency sounds and do not appear in the two charts mentioned because the curves were smoothed for purposes of reproduction.

Figure 5 shows the harmonics of backgrounds of 75, 250, and 500 cycles in some detail, together with the masking produced by these tones.[13] The dashed segments of the masking curves define the regions in which the signal was detected by means of beats. The height of the vertical line at the fundamental or masking tone frequency indicates the sensation level of the background. The heights of the vertical lines at the harmonic frequencies indicate the sensation levels of the signals which gave the strongest beats. Since the strongest beats occur when objectively presented beating tones have equal amplitudes, the heights of the vertical lines are assumed to indicate the sensation level of the aural harmonic.

This line of reasoning is indirect, but its results are in agreement with direct measurements of harmonics in the cochlear potentials, as described below. The agreement between these two sources of evidence indicates the validity of the best beat method, which equates stimulation and masking, and thereby supports the assumption used in constructing Figures 3 and 4.

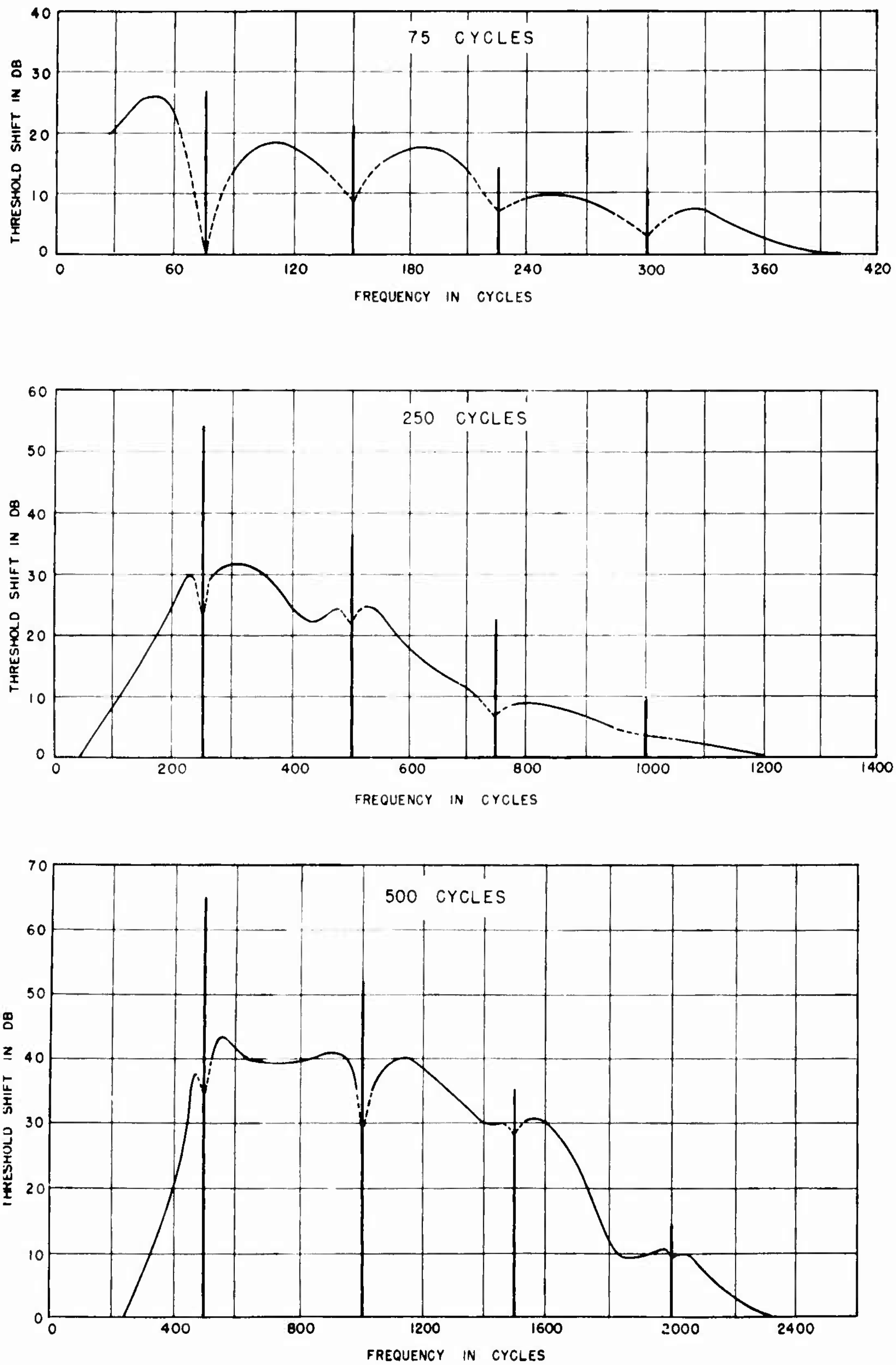


FIGURE 5. Masking of tones by tones (75, 250, and 500 cycles). (Courtesy *Journal of the Acoustical Society of America.*)

The results of analysis of aural harmonics by means of the method of beats are summarized in Figures 6, 7, and 8. These figures are based on a study of tones with frequencies between 50 cycles and 8 kilocycles, and with intensities between 60 and 140 decibels, above 2×10^{-4} dyne per square centimeter.

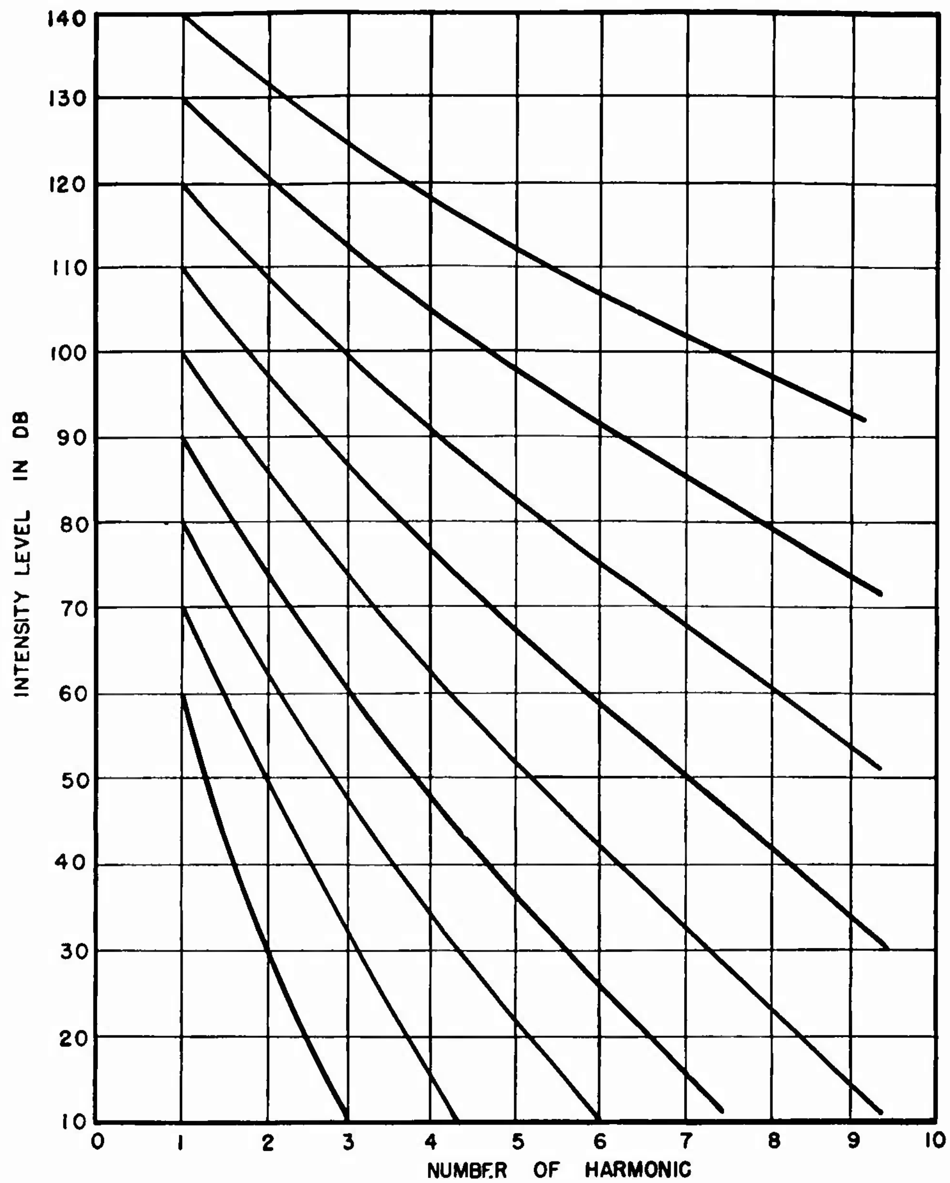


FIGURE 6. Relative intensities of aural harmonics produced by fundamentals of different intensities. (Courtesy *Journal of the Acoustical Society of America.*)

Figure 6 shows the number of the harmonic on the horizontal scale, where the first harmonic represents the fundamental (or objective stimulus frequency), and the intensity level of the fundamental and its various harmonics is given by the vertical scale. It will be seen that, for all the frequencies studied, the magnitudes of successive harmonics depend only on the intensity of the fundamental and not on its sensation level or frequency. Since the harmonics are due to the nonlinear nature of the ear's force-displacement characteristic, this result indicates that the amount of overloading depends predominantly on the incident pressure and not on frequency.

This figure also shows that the sensation levels of the harmonics may sometimes exceed

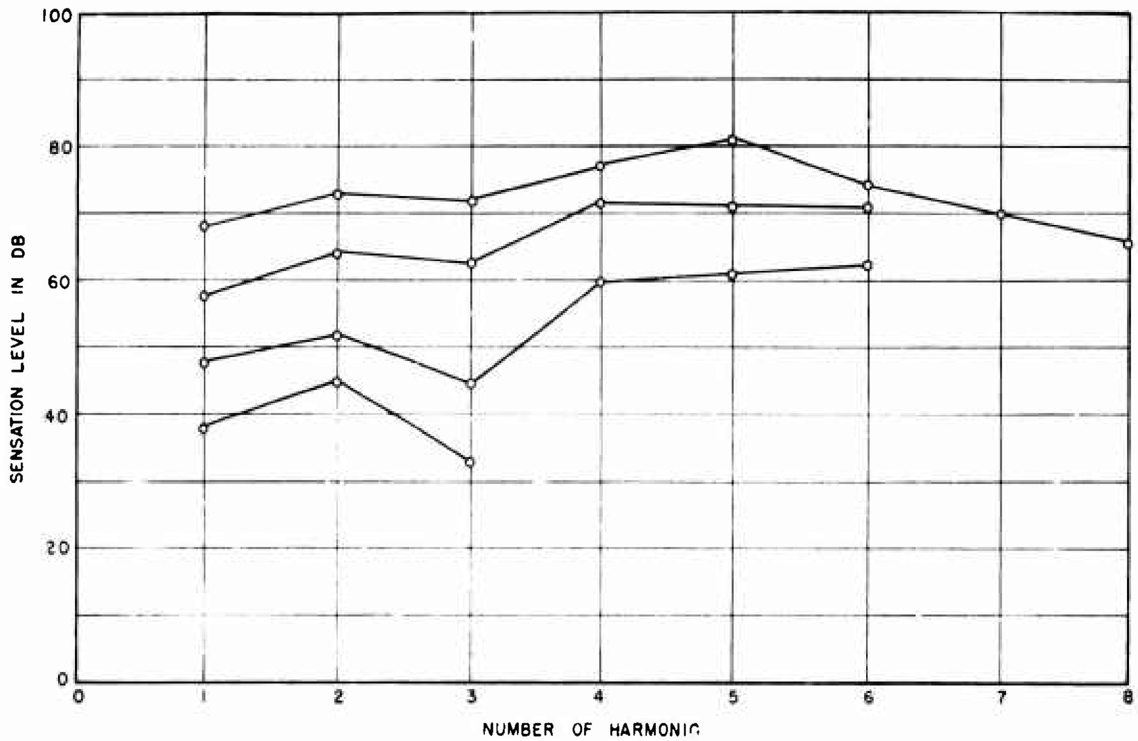


FIGURE 7. Sensation levels of harmonics produced by a 50-cycle tone. (Courtesy *Journal of the Acoustical Society of America*.)

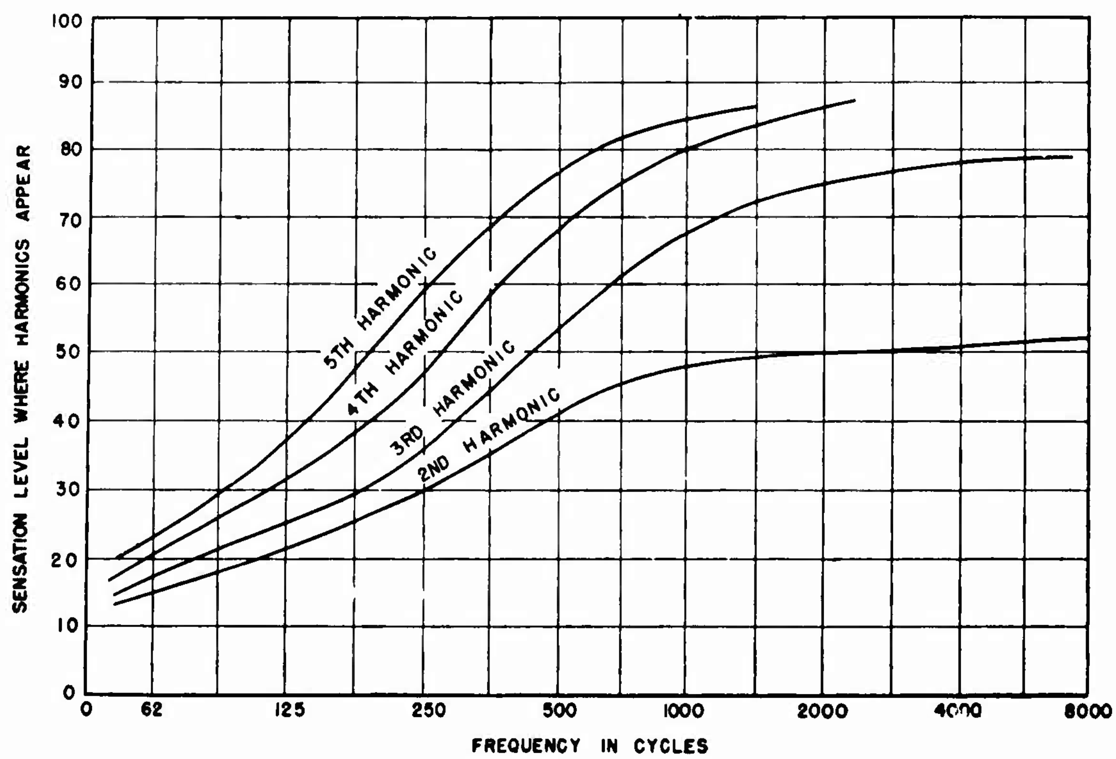


FIGURE 8. Sensation levels required to produce aural harmonics. (Courtesy D. Van Nostrand Co.)

that of the fundamental. Consider a 50-cycle tone at a sensation level of 70 decibels. From Figure 1, such a tone has an intensity level of 140 decibels. By comparing Figures 1 and 6 it will be seen that the intensities of the first four harmonics of a 140-decibel fundamental decrease at a rate of about 10 decibels per octave, whereas the audibility threshold between 50 and 200 cycles falls off at a rate of about 15 decibels per octave. In other words, the sensation levels of the harmonics increase at a rate of about 5 decibels per octave.

The sensation levels of the harmonics of a 50-cycle tone have been computed in the manner just outlined, and are shown in Figure 7. From this figure, it is clear that, for the stronger stimuli, the first six overtones have higher sensation levels than the fundamental. This explains the observation[13] that the first harmonic of a 50-cycle tone could always be sensed, no matter how faint the stimulus tone sounded. It also explains why the subjective loudness of low-frequency sounds increases so much more rapidly with increase of sensation level than does the sensed loudness of higher-frequency tones (see Figure 11). From Figure 8 it will be seen that the ease with which a tone produces aural harmonics is nearly constant above 1 kilocycle and increases rapidly for frequencies below this value.

2.2 DISCRIMINATION FOR TONES

The ear's ability to discriminate between tones of different pitch and loudness helps to determine the effectiveness of the ear in distinguishing a signal from the background. A study of the ear's performance for tones has revealed certain simple relationships which are of basic importance in psychoacoustics. Since these relationships are helpful in understanding the results given in following chapters, they are stated briefly in the present section.

2.2.1 Pitch

According to the theory of hearing presented in Chapter 1, sounds of different frequency set into vibration, or stimulate, different regions of the basilar membrane. The ability of the brain to distinguish between sounds received in different regions of the membrane then gives rise to perceived pitch. As noted in Chapter 1, this association between pitch and the position of maximum stimulation on the basilar membrane is called the place theory.

Equation (5) of Chapter 1 provides an approximate relationship between the fractional distance along the basilar membrane and the frequency of sound which produces maximum stimulation at that position. The values found from this equation are shown in Figure 9. The horizontal scale indicates the frequency of the stimulus tone, plotted logarithmically. The ordinate specifies the fractional distance (namely, x, or $s/3.1$) between the apex and base of the membrane at which the frequency in question produces maximal stimulation. The points represent the results of measurements obtained by techniques described in succeeding paragraphs. As seen, the measurements are in satisfactory agreement with the curve plotted from equation (5) of Chapter 1.

Furthermore, the figure indicates that over the range between 0.5 and 10 kilocycles, the position stimulated is approximately proportional to the logarithm of the stimulus frequency. Because ability to distinguish differences of pitch is a linear function of position on the basilar membrane, the common practice, which is followed in succeeding chapters, of plotting frequencies to a logarithmic scale is justified, since it provides a visual emphasis which approximates the auditory emphasis.

The filled-in circles plotted in Figure 9 were obtained in post-mortem studies of deafened human ears and show the relation between the position of lesions among the hair cells and the frequency region in which the individual's hearing was subnormal. This is the most direct line of evidence.

The results of indirect evidence from the measured functional characteristics of human ears are shown by the squares obtained from studies of frequency discrimination and described later in this section, and by the triangles which represent studies of pitch bisection. The term "pitch bisection" refers to the operation of finding a frequency which produces the sensation having one-half the pitch evoked by a comparison frequency. Various observers agree fairly well in their selections of the frequency which has associated with it a pitch sensation corresponding to half that produced by a standard, provided that previously learned relationships between musical intervals are disregarded.

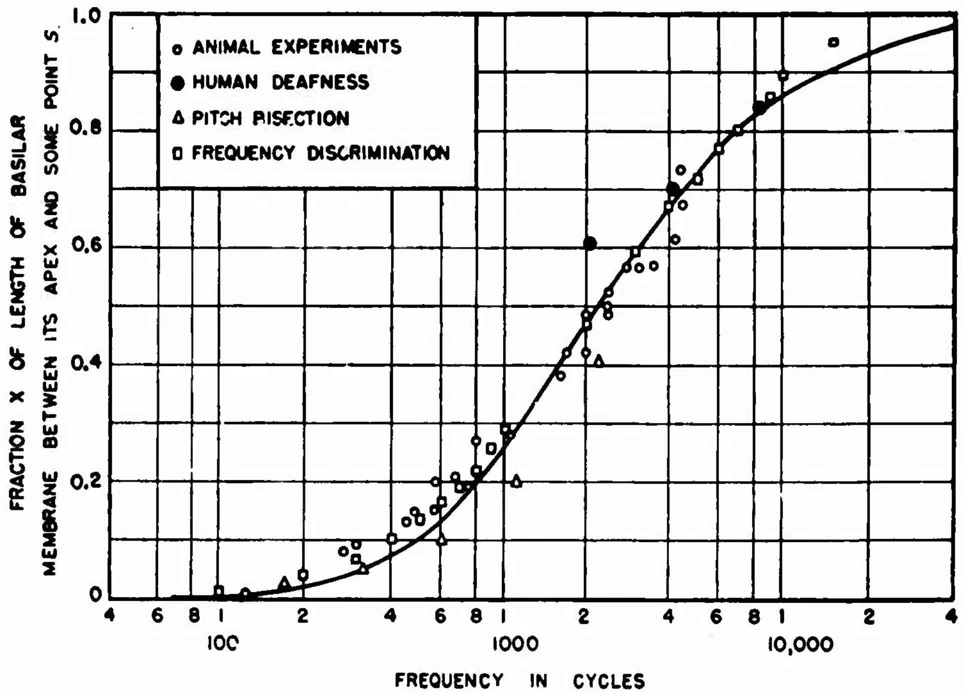


FIGURE 9. Position of maximum stimulation produced by different frequencies.

The open circles in Figure 9 represent measurements made upon experimental animals with hearing mechanisms very similar to that of man. Since the lengths of the basilar membranes in such animals usually differ from that in human ears, the measurements of position have been expressed as fractions of the total length of the membrane, so that the same vertical scale would apply in all cases.

Animal measurements have been made in several ways. One method consists of finding the position upon the outer surface of the intact cochlea at which the maximum value of the cochlear potential occurs for an impressed tone of given frequency and of repeating the measurement for tones of different frequencies. About half of the circles in Figure 9 were obtained by means of this technique.[14,15]

By carefully drilling into the cochlea and producing localized injury of the inner ear structure, the remainder of the open circles in Figure 9 were obtained. In order to correlate the position of the injury with frequency, the magnitude of cochlear potentials produced by various tones were measured before and immediately after the operation. In general, localized injury resulted in diminished response to tones in a fairly narrow frequency region. In Figure 9, the midpoints of these regions have been plotted against the corresponding position coordinates of the injuries.

Indirect evidence confirming the validity of Figure 9 for human ears has been obtained by observations on the performance of the ear under different conditions. The general consistency of Figure 9 with the observations on masking is shown in Figure 17.

Although Figure 9 gives the position of stimulation of the basilar membrane for sound of a given frequency, it does not indicate directly the least frequency difference which the ear can detect. This minimum perceptible frequency change is called the *frequency limen.* This frequency sensitivity of the ear would be expected to depend both on the width of the stimulated region on the membrane and on the nature of the processes by which stimulation of the membrane is converted into a nerve impulse and transmitted to the brain. It has been shown in Chapter 1 that the vibration of the strings in the basilar membrane is transmitted to the auditory nerve by means of the hair cells. It may be anticipated that no difference in frequency between two successive tones can be detected if it corresponds to a shift of position along the membrane by less than the spacing between adjacent groups of cells, that is, by less than the spacing between the Corti arches. Since there are about 6,000 of these members in the ear, the frequency limen may be expected to correspond to a change of x by at least $1/6{,}000$ or 1.67×10^{-4}. If the stimulation of the membrane is distributed over a very wide region, even greater differences in frequency might go undetected however.

The observed values[16] for the frequency limen δf are given by the plotted points in Figure 10. In these tests, the frequency of a sustained tone was shifted up and down, about 3 times a second, over a range Δf cycles. The values of Δf for which the warbling effect could just be distinguished are plotted in the figure. With somewhat different methods of presentation about the same values were obtained,[16] although with faster or slower frequency sweeps the ear became less sensitive. The sensation level of the tone used was about 40 decibels; the results do not depend critically on the sensation level for higher levels. It is necessary, of course, that the two tones to be differentiated be presented successively. If presented simultaneously, even a slight frequency difference would give rise to beats which could readily be heard.

The data in Figure 10 permit the computation of the smallest number of frequency steps between the lowest and highest audible frequencies (about 1,500). Assuming that each step requires a shift of the point of maximum stimulation on the basilar membrane by a fixed distance gives the squares shown in Figure 9.

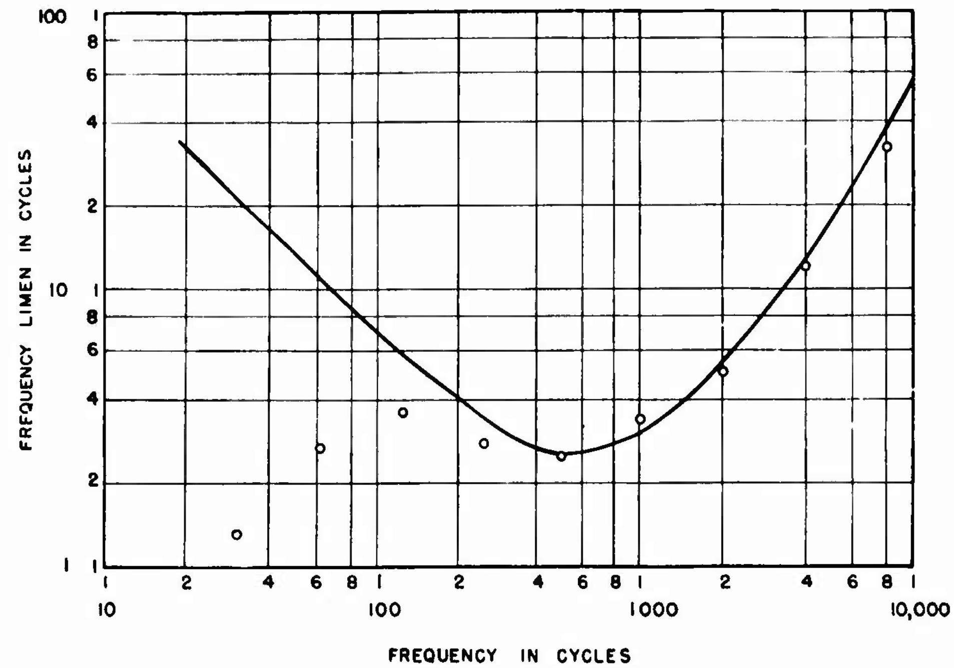


FIGURE 10. Frequency limen for different frequencies at a sensation level of 40 decibels.

The smooth curve drawn in Figure 10 was obtained by differentiating equation (5) in Chapter 1. It will be seen that the agreement between the theory and the observations is good for frequencies above 400 cycles per second. Disagreement for the lower frequencies is believed[16] to arise from the aural harmonics which these tones so readily generate.

Figure 10 indicates that the stimulated region of the membrane is relatively narrow when a single tone is present. However, the evidence on the masking of tones by tones, presented in the previous section, suggests that this region is relatively wide. The data on the masking of tones by wide-band sounds, discussed in Section 2.3, indicate that a pure tone sets into vibration a segment of the basilar membrane whose relative width Δx is about 20 times that used in computing the curve shown in Figure 10. It is not clear how the ear can determine so accurately the center of the excited region. Various suggestions along this line are discussed in reference 17. Although the theoretical basis of Figure 10 is not wholly clear, this figure gives a reliable indication of the frequency sensitivity of the ear under ideal conditions.

2.2.2 Loudness

The subjective loudness of a sound is apparently closely related to the masking properties of the sound and to the detectability of the sound in the presence of masking backgrounds. In addition, the ability to detect one sound in the presence of another depends on the loudness of each sound separately. Ability to detect small changes of loudness under various conditions is of crucial importance in listening for modulated sounds. Thus a study of loudness and loudness discrimination is important to the specific subject of underwater sound detection as well as to psychoacoustics in general. The meaning of loudness is discussed in the following paragraphs, while loudness discrimination is treated at the end of this section.

Measurement of Loudness

Since the ear is not a quantitative measuring instrument, it is not possible to measure directly the subjective loudness of a sound. The observer can, at most, state that one sound has greater loudness, the same loudness, or less loudness than another sound. Since the ear can identify two sounds as having the same loudness, it is at least possible to adjust the relative intensities of widely different sounds until they have the same loudness. Both systems of measuring loudness which are now in use start from this basic fact. Fortunately most observers with normal hearing usually agree with each other as to whether or not one sound is equal in loudness to another.

The simplest measurement of loudness is to refer all sounds to a sustained tone of 1,000 cycles per second. The loudness of any sound may then be measured in terms of the intensity level of the 1-kilocycle tone of equal loudness; this is equal to the sensation level of the 1-kilocycle tone (see page 17). The level so determined is called the *loudness level* of the sound. The phon is used for giving the loudness level. A sound whose loudness is the same as that of a 1-kilocycle tone at an intensity level of 50 decibels has a loudness of 50 phons. In general, a loudness of 70 phons is regarded as a comfortable listening level for typical sounds. Sustained tones of high pitch at this level can become very annoying.

Two tones of different frequencies but of the same loudness level will usually have different intensity levels. All tones at the threshold of audibility have zero loudness. It is not true, however, that all tones at the pain threshold are equally loud.

Equal loudness contours are shown in Figure 11, where the intensity levels for sounds of constant loudness are plotted against frequency. The thresholds of audibility and pain correspond, of course, to those shown in Figure 1. Since the contours approach each other at very low frequencies, it is evident that given changes of loudness correspond to smaller intensity changes at the very low frequencies. This effect is presumably owing to the larger stimulation contributed by aural harmonics of low-frequency tones (see Figure 8).

Another method of measuring loudness has been developed which presumably has more objective significance. In this method, it is assumed that the loudness of a sound heard with two ears is equal to twice the loudness of the same sound heard with only one ear. On the basis of this assumption, the observed data can be used to draw a curve for each frequency connecting the loudness with the intensity level. Other similar assumptions lead to much the same results, and there is some reason to believe that these methods do in fact measure subjective loudness. Since these results have not as yet been applied to the study of the recognition of underwater sounds, they are only mentioned in passing here.

Loudness of Pulses

The data summarized in Figure 11 refer to the loudness of sustained tones. Some work has also been carried out by the Bell Telephone Laboratories on the perceived loudness of short pulses of varying duration. These results have been made available in advance of publication (see also reference 22), and are reported here because of their relevance to the masking of short pulses by noise and reverberation.

In these tests the intensity level of a short pulse was varied until its subjective loudness was the same as that of a sustained tone at the same frequency and of known loudness level. A value was then read from Figure 11 for the loudness which the sustained tone would have if its intensity were just equal to that of the pulse. The measured loudness level of the pulse was then plotted against the computed loudness level of the hypothetical tone of equal intensity. In this way, for each of the five pulse lengths used a curve was drawn showing the loudness level of the pulse as a function of the loudness level of the equally intense sustained tone of the same frequency. The tests were carried out for three frequencies, 125, 1,000, and 5,650 cycles per second. The resulting plots are given in Figures 12, 13, and 14.

It is evident from these figures that when the loudness level of the tone is appreciable, a reduction in the duration of the presented tone results in a marked decrease of the subjective loudness. This is a result of what has been called the finite build-up time of the ear. Other evidence also indicates that when the duration of a sound is less than ¼ second, the effectiveness of aural perception diminishes, and the subjective loudness decreases.

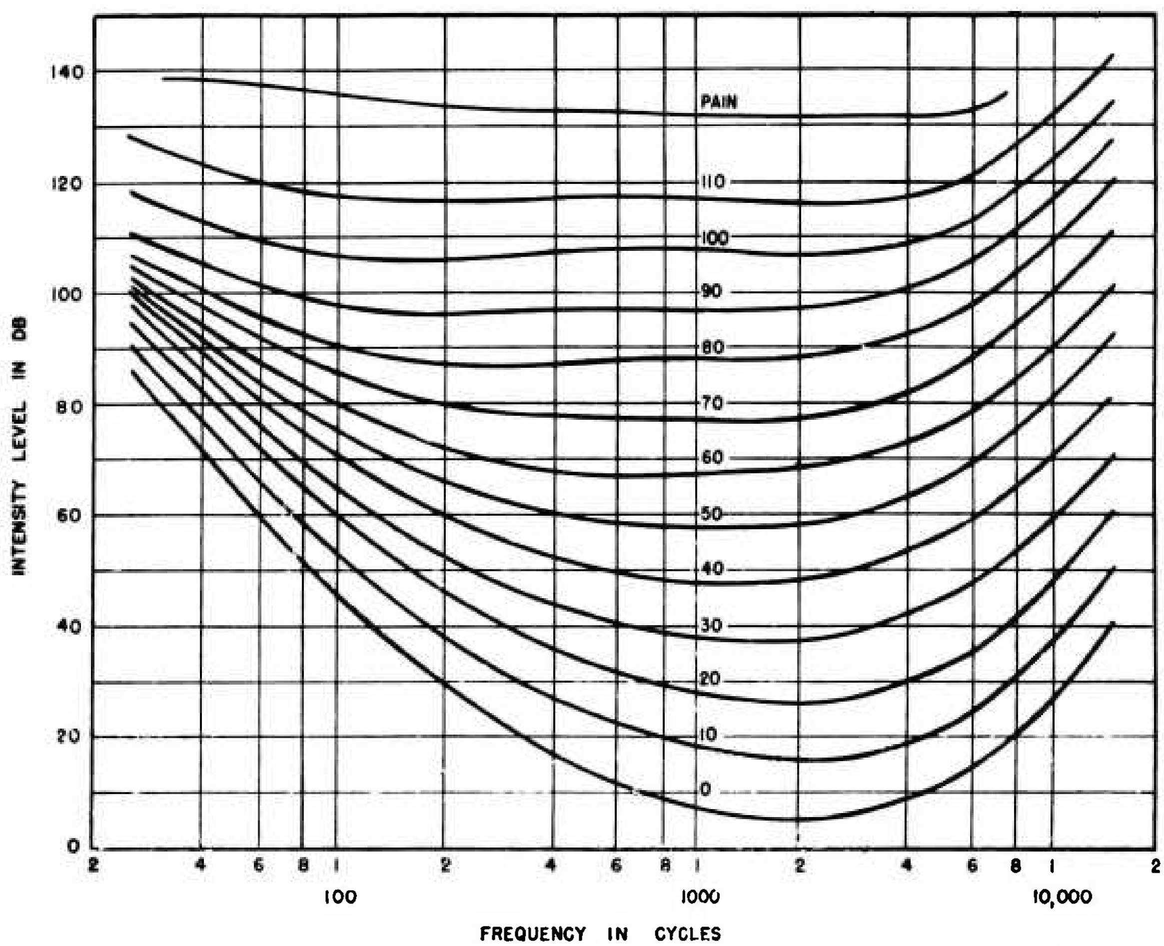


Figure 11. Equal-loudness contours. Loudness levels in phons are indicated on curves. (Courtesy *Journal of the Acoustical Society of America.*)

Figures 12 through 14 seem to show, however, that as the auditory threshold is approached, the finite build-up time of the ear seems to lose its effect. A tone which is barely audible when sustained, in the absence of any masking background, can apparently still be heard when its duration is reduced to as little as 0.05 second. While the curves are obviously extrapolated for very low loudness levels, the general nature of this extrapolation seems well indicated by the data. This result is presumably of considerable psychoacoustic significance in that it casts important light on the properties of the ear's build-up time.

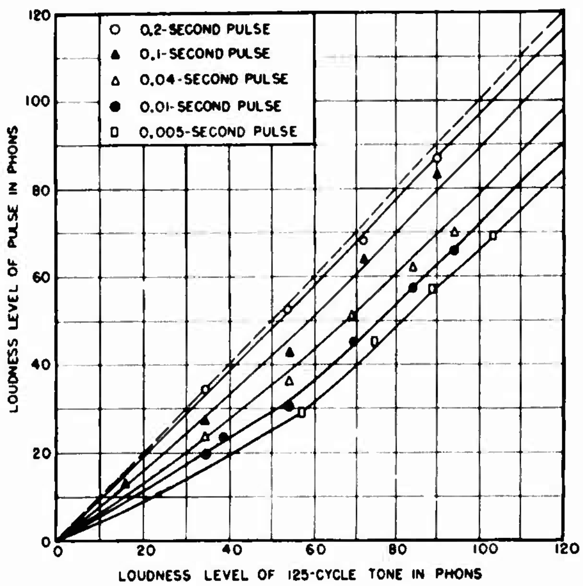


FIGURE 12. Loudness levels of tonal pulses (125 cycles). (Courtesy Bell Telephone Laboratories.)

Loudness Discrimination

When two tones with slightly different loudness are presented successively, or when the intensity of a tone is varied, the change of loudness can be detected only if it is greater than a certain value. The minimum perceptible intensity increment, which depends on the rate of modulation, the sound frequency, and the loudness level, is called the *intensity limen*. For young observers with normal hearing, measured values[18] of the intensity limen are shown in Figure 15 for all values of the frequency shown on the horizontal scale. The rate of amplitude modulation used in these tests amounted to 3 cycles per second. The vertical scale at the right gives the minimum intensity change in

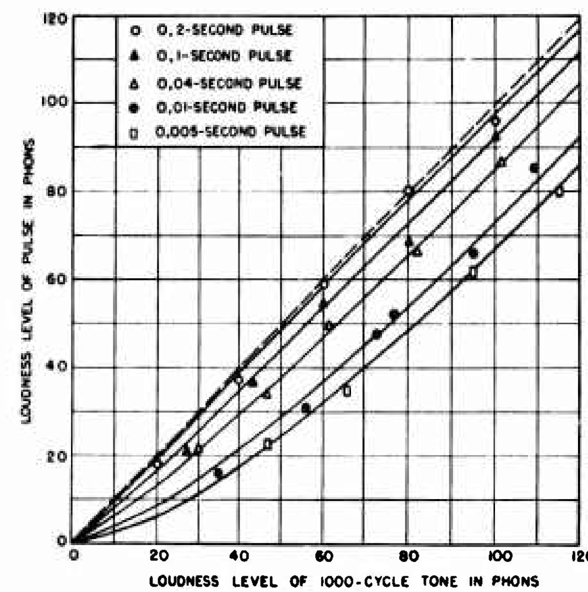


FIGURE 13. Loudness levels of tonal pulses (1,000 cycles). (Courtesy Bell Telephone Laboratories.)

decibels, while the scale on the left gives directly the relative intensity change $\Delta I/I$.

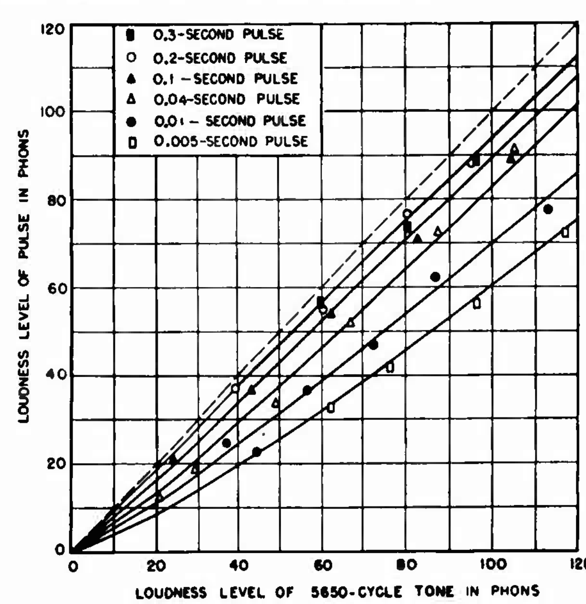


FIGURE 14. Loudness levels of tonal pulses (5,650 cycles). (Courtesy Bell Telephone Laboratories.)

Figure 16 shows the dependence of the intensity limen on the rate of amplitude modulation of a 1-kilocycle tone, for loudness levels of 25 and 50 decibels. It indicates that intensity discrimination improves with intensity. About the same results would probably be obtained at frequencies between 0.6 and 10 kilocycles. For example, results similar to those in Figure 15 were obtained in a study with a wide band of thermal noise.[19] of whatever acoustic character. Hence it is usually best to distinguish between background, or unwanted sounds, and signals, or wanted sounds.

When a complex sound, with components distributed over a large band of frequencies, is incident upon the ear, the latter continues to act as an analyzer. Thus, it is possible to focus attention upon different pitch regions in such a sound, and the effect of removing some of its

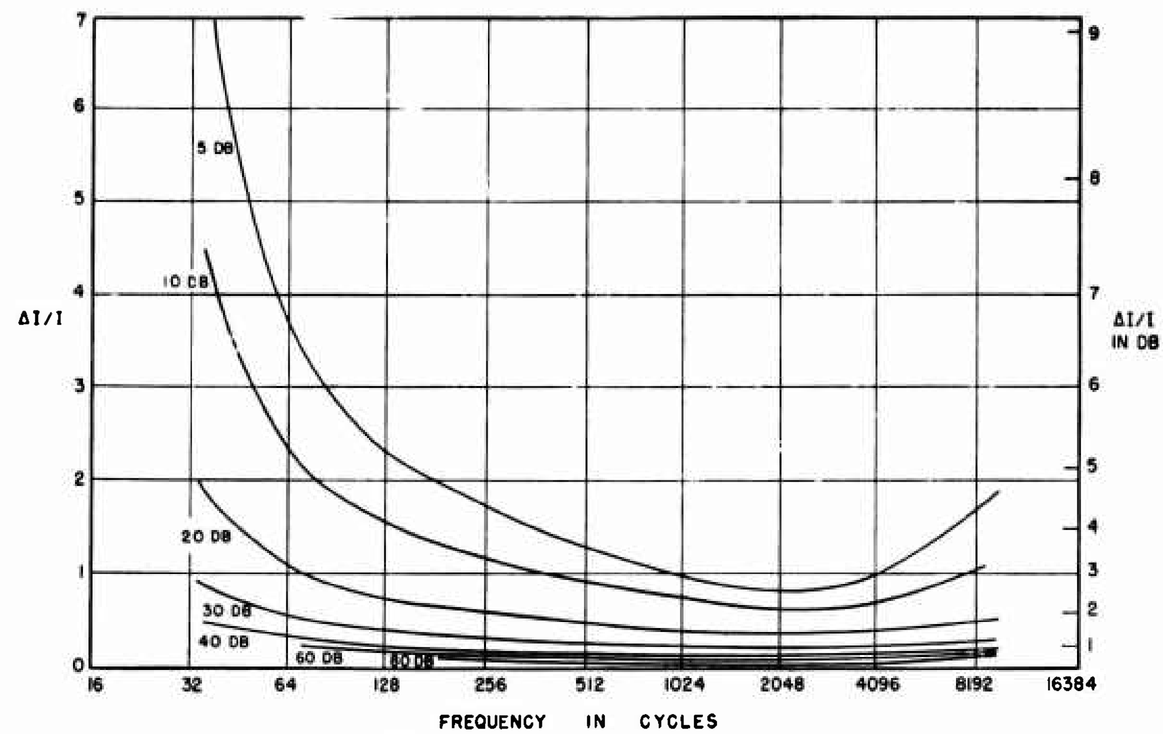


FIGURE 15. Intensity limen at different frequency and sensation levels. (Courtesy *Physical Review*.)

2.3 COMPLEX SOUNDS

By a complex sound is meant a pressure disturbance with a complex acoustic spectrum. Thus, at one extreme is a pure tone, corresponding to a line spectrum. An intermediate case is illustrated by a musical chord, or even by the sound obtained when all the keys of a piano keyboard are struck simultaneously, giving a very complicated line spectrum. At the other extreme is the type of sound produced by breaking surf, containing all possible frequencies within the sonic band.

The term noise is often used to denote a complex sound, and many, but not all, common noises are of this character. However, the term noise is also used to mean an unwanted sound

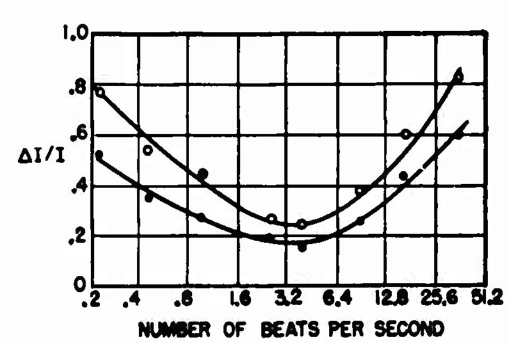


FIGURE 16. Dependence of the intensity limen on the rate of variation in the intensity of a 1-kilocycle tone. The open circles refer to a sensation level of 25 decibels and the filled-in circles to a sensation level of 50 decibels. (Courtesy *Physical Review*.)

frequencies by electrical or acoustic filtering is quite perceptible.

The masking and loudness produced by complex sounds are in general simpler to predict than are the corresponding properties for individual tones or groups of tones. In the case of a single tone, a large part of the incident energy goes to stimulate a region on the basilar membrane tuned to the stimulus frequency, but some portion of the incident energy is also scattered to other frequencies because of the nonlinear character of the ear's response. The resultant masking and stimulation is therefore strongly dependent on the harmonics, as well as on combination and difference frequencies, introduced by the hearing mechanism.

In the case of complex, or wide-band sounds, the scattering of energy to portions of the basilar membrane not tuned to the stimulus frequency occurs for each of the components in the stimulus. In any one region about as much sound will usually be scattered out as is scattered in. Thus the net result of this process of multiple scattering is that the stimulation at any point on the membrane is determined essentially by the fraction of the energy in the complex sound contained in a narrow frequency band, called a critical band, centered at the frequency in question. In other words if the distribution of energy among the various frequencies in the objective sound is not so skewed that the aural harmonics generated by the low-frequency content have higher sensation levels than are produced by the objective components at those frequencies, then all the masking of any consequence will be adjacent masking. As

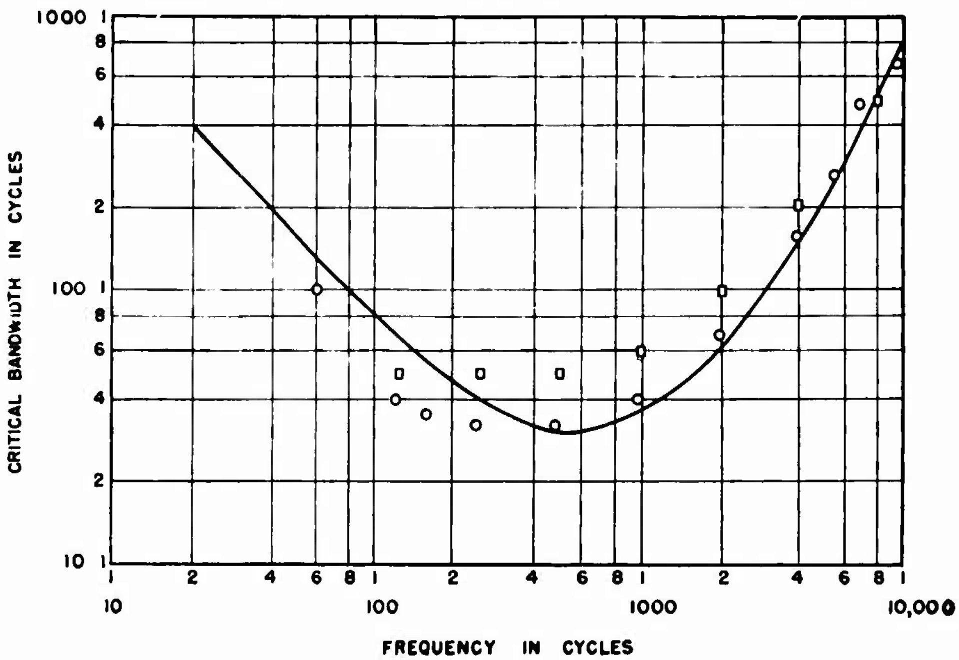


FIGURE 17. Aural critical bands for signal-background mixtures presented to both ears.

Figure 6 indicates, the levels of harmonics fall off quite rapidly even when the fundamental is at the pain threshold. For ordinary listening, therefore, it is not likely that complex backgrounds will produce remote masking unless the energy of the various components decreases more rapidly with increasing frequency than about 20 to 30 decibels per octave.

Just as all the masking produced by complex sounds is essentially adjacent masking, so also the stimulation and loudness they produce can be computed directly from the objective spectra, without regard to the harmonic content introduced by the ear, unless the complex sounds contain salient peaks or drop off at rates in excess of 20 decibels per octave. Subject to this restriction, the loudness of a complex sound is generally the sum over all critical bands of the loudness in each band.

The facts concerning masking by distributed sounds have been determined in the following way.[20,21] A wide-band sound (thermal noise in the case studied), with a flat spectrum extending from, say, 400 to 1,200 cycles, is presented to the ear as a constant-level background. Then the level of an 800-cycle signal which is just audible in the presence of this background is determined. Successive redeterminations of this signal level are made, keeping all the conditions the same as already described, with the exception that a band-pass filter centered at 800 cycles is introduced to eliminate some of the background components.

Since, as stated before, remote masking is not very significant in the case of wide-band sounds, the level of the just detectable 800-cycle signal is not affected by removing the upper and lower frequencies from the background band. This constancy of the level of the just detectable 800-cycle signal continues to be observed as the band of admitted background frequencies is narrowed, until a critical width of the filter band is reached. For an 800-cycle tone masked by distributed noise this critical band width amounts to 40 cycles (see Figure 17). For narrower bands of noise (centered at 800 cycles), the level of the just audible 800-cycle tone diminishes in direct proportion to the decrease in the band width of the background; that is, when the band is cut from 40 to 20 cycles, the intensity of the just audible signal drops by a factor of 2.

Similar experiments, conducted for tones of other frequencies, give the widths of critical bands centered at different points in the sonic range of frequencies, and these are also shown by the plotted points in Figure 17. Thus, in the presence of distributed sounds, the ear behaves as though it were provided with a group of band-pass filters which permit it to eliminate masking interference from all components beyond the cut-off limits of the filters. The tests show, furthermore, that a sustained tonal signal is just audible when its intensity is equal to the intensity of the background components contained in the critical band centered at the *tone* frequency.

Evidence discussed in Chapter 4 indicates that the critical band criterion also applies when the signal is a distributed sound. Under these circumstances, the signal becomes detectable in the presence of a distributed background when the intensity of the signal equals the intensity of the background in at least one critical band. However, fluctuations of signal and background level may modify the latter result in certain cases (see Figures 29, 38, and 40 in Chapter 4).

The width of a critical band is determined essentially by the fact that even when the ear is stimulated by a pure tone of low intensity the basilar membrane vibrates in patches of width Δs. This width has been estimated in Section 2.1.2 as being about 3×10^{-2} centimeter, or 1 per cent of the total extent of the membrane. This same result may be obtained from the critical band width Δf given in Figure 17 and the curve in Figure 9. By finding the value of Δx on the ordinate of Figure 9 which corresponds to the value of Δf on the ordinate of Figure 17, for a given value of the frequency, it will be seen that the value of Δx corresponding to one critical band is very nearly equal to 0.01, that is, 1 per cent, for all frequencies. Since $x = s/3.1$ (see Section 1.2), $\Delta x = \Delta s/3.1$; hence, when Δx is 0.01, Δs amounts to 3×10^{-2} centimeter.

With this value of 0.01 for Δx, values of Δf have been computed. The computation is elementary, since Δf should equal df/dx times Δx, where df/dx may be obtained either by differentiating equation (5) of Chapter 1 or from the graph in Figure 9. The values found in this way are shown by the smooth curve in Figure 17. It is evident that the agreement is close. This figure may be compared with Figure 10, showing the corresponding agreement between predicted and theoretical values of the frequency limen. This comparison shows that the critical band width Δf at each frequency is about 20 times the frequency limen at the same frequency. This general agreement between observation and theory strengthens confidence in the concepts developed in this chapter and the preceding one and indicates that these concepts can be reliably applied to a tentative interpretation of practical data. However, it is evident that other factors, such as the nature of nerve conduction and the build-up time of the ear, are not as yet wholly understood and may be expected to affect practical results in unforeseen ways.

Chapter 3

CHARACTERISTICS OF TARGET SOUNDS AND NOISE BACKGROUND

THE STUDY of underwater sound detection, or recognition, may be divided into two parts, depending on the type of sound equipment and how it is used. Sound gear may be used to detect the sound radiated by some object and is then called listening gear. It may also be employed to detect sounds first radiated by the gear and then reflected back from some object; such installations are described as echo-ranging gear. The object which either generates or reflects the desired sound is often called the target, and the signal which it directs toward the gear is known as either the target sound or the echo, depending on its origin. Obviously, echo-ranging gear can be used as listening gear, and in current practice the operator alternates both these uses. Chapters 4 through 6 discuss the recognition of target sounds, while Chapters 7 through 11 discuss the recognition of echoes.

The hydrophone which receives the acoustic energy is usually an electromechanical converter which transforms the sound impulses into equivalent electric impulses. These electric impulses are subsequently amplified and may be changed in other ways which affect overall performance. Some of these changes will be discussed in this chapter. The output of the hydrophone and its associated circuits is finally coupled to the mechanism used to detect the signal; the ear, together with headphones or a loudspeaker, is the detector considered throughout this volume.

The major factors determining the possibility of signal detection are the characteristics of the received signal and background, the properties of the gear, and the limitations of the detector. These factors will be discussed in sequence in this chapter. Chapters 4 and 5 describe the data obtained by different groups on the recognition of target sounds.

The character of the sounds received by the gear and presented to the operator is affected by a number of factors, few of which are subject to control. The typical signal discussed in Chapters 4 and 5, acoustic radiation from a surface ship or a submarine, depends among other things on (1) the orientation of the target vessel relative to the listening vessel, (2) the type of target vessel, (3) its speed, and (4) the operating condition of its machinery. The intensity, frequency composition, and phase relations of the radiated energy may all be modified during transmission from source to receiver.[a]

The amount and character of the interfering background depend on (1) the nature of the sounds produced by other sources, such as waves and whitecaps; (2) how well those sounds are transmitted to the hydrophone and to the ear; and (3) the existence of nonacoustic disturbances, such as electric fields produced by motors, which may be transformed into sounds by the various parts of the listening gear. Airborne sounds not received through the listening gear may also form part of the interfering background, although the effect of these noises can usually be minimized by increasing the amplifier gain.

Many listening hydrophones are designed to respond efficiently only to sounds whose angles of incidence with the hydrophone face lie within a fairly narrow cone. The axis of this cone is called the *hydrophone axis.* This ability of hydrophones to respond preferentially to sounds incident on the hydrophone axis is termed *hydrophone directivity,* and hydrophones which possess this property are called *directional hydrophones.* Hydrophone directivity increases as the size of the hydrophone increases and as the frequency of the incident sounds increases. Hydrophone directivity serves two functions: (1) it diminishes the amount of interfering background transmitted to the operator, wherever such interference reaches the hydrophone from directions that are not on its axis; and (2) it furnishes a method for determining the relative bearing of the target. Additional discrimination of the

[a] See STR Division 6, Volume 8.

receiver against various frequencies present in the background may be obtained by designing the gear to transmit some frequencies present in the background and to suppress others. This procedure may be undesirable since it suppresses the same frequencies in the signal; thus, it may discriminate against precisely those signal frequencies which permit ready detection of the target. As shown in Chapters 1 and 2, the ear possesses its own frequency-discriminating mechanism.

In order to evaluate the effects upon the ear of these various factors, a more detailed description of the characteristic features of ship signals and interfering sounds is given in Sections 3.1 and 3.2. Those features of the listening gear which affect the recognition of target sounds are described in Section 3.3.

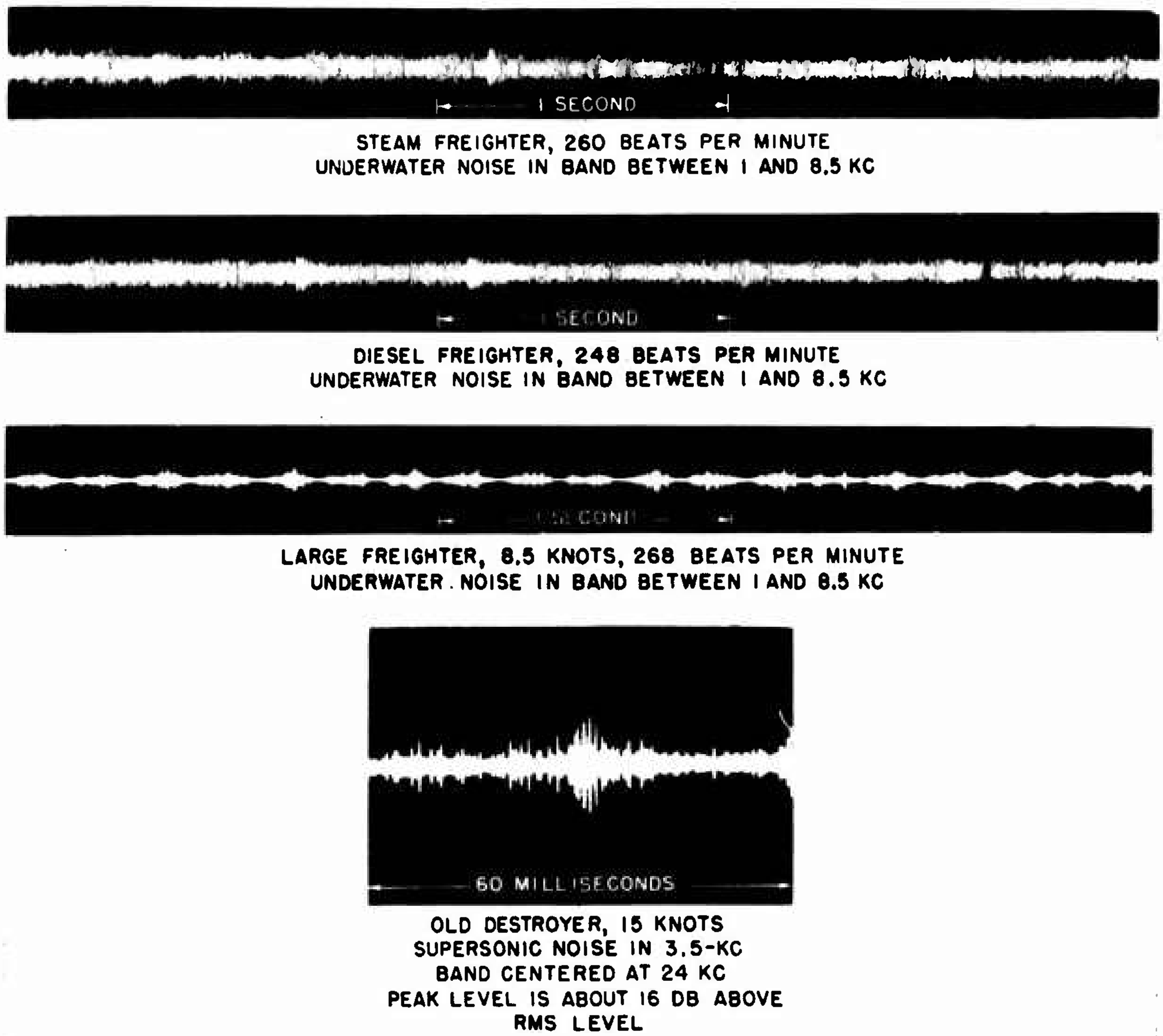


FIGURE 1. Oscillograms of sonic and supersonic ship noise. The playback of the signal was applied to the terminals of a nonpersistent CRO, and the resulting deflections were photographed on continuously moving film.

3.1 SOUNDS FROM SUBMARINES AND SURFACE VESSELS

The character of the acoustic energy radiated into the water by a submarine or surface vessel is modified during transmission.

The nature of the signal at the listening hydrophone, compared with that of the masking background, is what determines whether the target will be detected. Unfortunately, accurate measurements of this signal are difficult to make at distances comparable to the maximum distances at which listening gear can detect the source vessel, because the signal-to-noise ratio is quite unfavorable for most measuring devices at such long ranges. Consequently, the acoustic outputs of ships are usually measured at ranges of 50 to 500 yards, and the probable characteristics of the radiated sound at practical detection distances are deduced from fundamental studies on the transmission of underwater sound.

3.1.1 Close to Source

While ship sounds are often intense, they never represent more than a tiny fraction of the energy lost by the ship through various kinds of operating inefficiency. They are therefore not necessarily diminished by general improvements in design; specific attention to their causes is required in order to reduce their intensity. The radiated energy is acoustically complex; it is distributed throughout the spectrum and is appreciable even in the subsonic and supersonic regions. There are two major sources of sound: the propellers and the engine room.

Most of the sound coming from the propellers is produced by *cavitation*, that is, from the formation of cavities in the water near the propeller blades. These cavities are not due to the churning of air into the water; they are formed when the pressure behind the moving blades falls to a threshold value, which increases with increasing hydrostatic pressure. The rapid pressure changes which accompany the formation and collapse of these cavities constitute the radiated sound. When such cavities collapse against the blades, the impact is large; this effect probably accounts for the high rate of propeller erosion.

In practice, the amount of cavitation noise is determined by (1) propeller tip speed; (2) acceleration, including the application of helm; (3) loading, in the case of cargo vessels and transports; and (4) the depth of the propeller below the surface. By way of an example of the last point, submariners engaged in evasive maneuvers can diminish the amount of cavitation noise radiated by the submarine by operating at increased depth, although there is some speed, at any practical depth, which marks the onset of cavitation. For a given speed, cavitation noise is greater on the average for vessels of larger tonnage.

Cavitation sounds represent the total effect of a large number of independent events. Each of these events, the formation or collapse of a cavity, generates a brief pressure pulse with a complex acoustic spectrum. When cavitation noise is subjected to frequency analysis, it is found that the intensity within a narrow band is very nearly inversely proportional to the square of the frequency, for frequencies between 100 and 30,000 cycles per second.[b] In other words, the spectrum has a negative slope of 6 decibels per octave [$10 \log (f/2f)^2 = -6$ decibels]. The steep decline of intensity with frequency does not necessarily mean that the low frequencies are the most important from the standpoint of detectability, because detectability is determined by the *relative* strengths of signal and background at various frequencies. If desirable, the high frequencies may be amplified more than the low.

To the ear, cavitation is a rushing, boiling sort of sound without much character. What character cavitation sounds have is conferred by propeller modulation. Very often, all the frequencies in the observed cavitation sound are amplitude-modulated at the blade rate or the shaft rate or both. The percentage modulation which is observed depends on the type of vessel and the condition of its operation (see Figures 1 and 2), and is probably one of the factors which enables an experienced sound

[b] These frequencies are at present the practical upper and lower limits for most types of underwater sound gear; in a few applications, however, frequencies outside these limits are used. An upper limit is imposed by the increase in transmission loss with increasing frequency; a lower, by the increase in background interference with decreasing frequency as well as by the properties of the ear (see Section 2.1.1). Ship spectra do not continue to rise at the very low frequencies; there is often a peak in the neighborhood of 30 to 50 cycles per second and an abrupt decline for still lower frequencies.

operator to classify the target. Well-modulated propeller cavitation has been described as having a characteristic rasping sound.

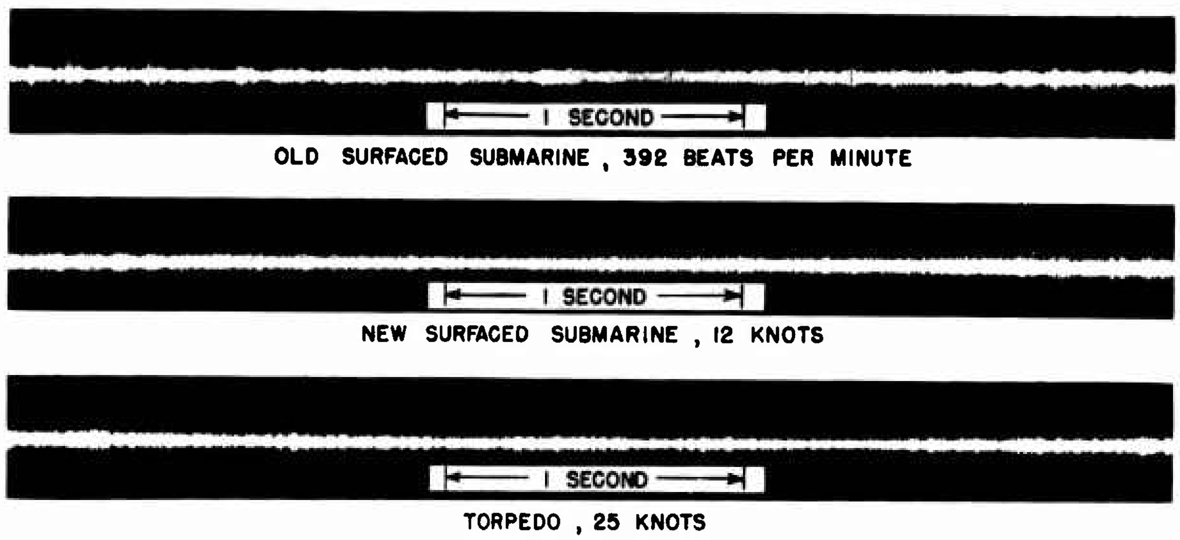


FIGURE 2. Oscillograms of submarine and torpedo noise in the band between 1 and 8.5 kilocycles. The playback of the signal was applied to the terminals of a nonpersistent CRO, and the resulting deflections were photographed on continuously moving film.

The blade and shaft rate modulations are often quite distinct in character, and when both are present the propeller sounds usually have an accented beat, so that every third or fourth "chug" is louder than the intervening ones. For multi-screw vessels, there are additional possibilities for complicated rhythm patterns. The accent of propeller sounds is often characteristic of a particular ship and provides an additional method of identifying the source. A count of the number of propeller beats per minute is useful in estimating the speed of a target; when the rate is too high for comfortable count of the individual beats, it is helpful to count the accented beats only, if the beat is accented, and to multiply by the number of beats in the rhythm cycle.

In addition to cavitation noise, propellers may produce more nearly tonal kinds of sound. Improperly designed propellers often show characteristic torsional or flexural vibrations which produce very intense tones. Figure 3 shows the spectrum of a ship with a "singing" propeller. The single-frequency peak, shown as a vertical line at 1,100 cycles, is about 35 decibels higher than the spectrum level of distributed sounds at neighboring frequencies. The remainder of the spectrum is typical of cavitation, and has a slope of about 6 decibels per octave. The 1,100-cycle tone observed in

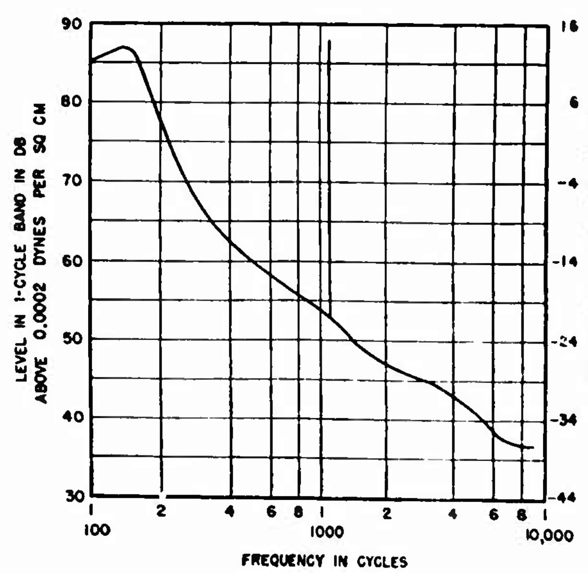


FIGURE 3. Spectrum of an aircraft carrier at 15 knots, measured 220 yards from the source.

this case was heard as a whine which was amplitude-modulated at the propeller rate, showing that it originated at the screws.

Finally, propeller shafts occasionally produce an assortment of characteristic moans, squeaks, and howls. These sounds are particularly likely to occur as the result of damage sustained in action, and the increased noise constitutes an additional operating hazard until repairs can be made.

Obtaining the spectrum of a noise source consists, essentially, of measuring the average energy contained in narrow frequency bands at various frequencies. As shown in Chapter 2, adequate resolution of the spectrum (from the point of view of loudness and masking) may be obtained by analyzing the sound with band-pass filters about 50 cycles wide. When a distributed sound is measured in this way, the total energy passed by the filter increases with the width of its pass band. If a pure tone is measured through a filter which passes the tone frequency, the measured energy is independent of filter width. Therefore, the intensity of a pure tone mixed with a distributed sound may be computed by noting the difference between the reading obtained when the filter passes the tone together with the distributed sound, and the reading obtained when the filter passes only the distributed sound at frequencies in the immediate neighborhood of the tone. Alternatively, two filters of different widths may be used to pass the mixture, and the contribution due to the tone may be estimated by the method described in Section 4.1.3.

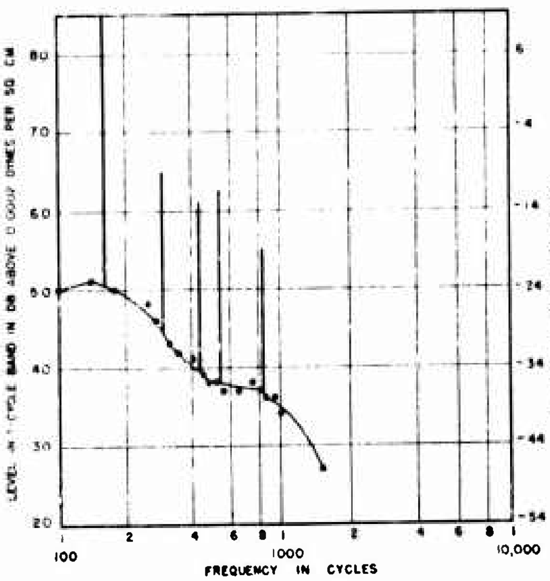


FIGURE 4. Spectrum of a fleet-type submarine at 50 rpm (2.5 knots), and periscope depth, measured 150 yards from the screws.

Sounds entering the water from the interior

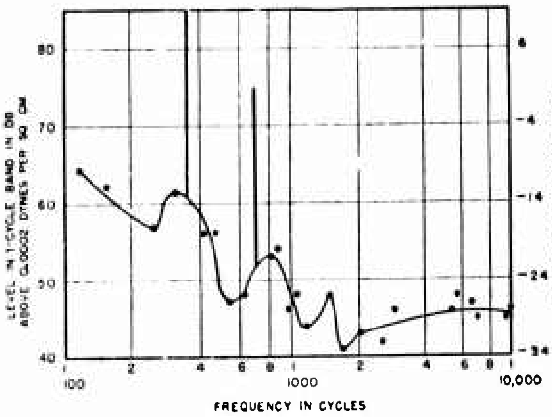


FIGURE 5. Spectrum of a fleet-type submarine at 120 rpm (6 knots), and periscope depth, measured 170 yards from the screws.

of the vessel generally originate at either the propulsion or the auxiliary machinery. Some of these sounds increase in intensity with increasing speed; others are independent of speed and may persist even when the vessel is adrift. Where such machinery is not mounted on an insulating base, most of the sound reaches the water by conduction through the supporting structures. However, even when vibration mounts are used, some energy reaches the hull by air conduction, and the transmission of this fraction can be reduced only by soundproofing the walls.

The spectra of machinery sounds have received little study, but several general remarks appear warranted. Such spectra may exhibit a group of fairly intense, harmonically related tonal components. This property is characteristic of many kinds of rotating machinery, such as reduction gears, ventilating fans, and electric generators. Such tonal components are about equally strong at frequencies below 1 kilocycle, diminish in intensity at higher frequencies, and are either very weak or relatively rare at frequencies above 3 kilocycles. Figures 4 and 5 illustrate this type of spectrum. Such tones often sound like the whine of a buzz saw and are usually frequency-modulated and amplitude-modulated. Occasionally, as in the case of reduction gear noise produced by submarines, they increase in frequency with increasing speed. Another type of machinery spectrum, heard as a grinding or rumbling sound, is illustrated in Figure 6. In general, machinery spectra appear to be rich in low-frequency components; they tend to be relatively flat in the region below 1 to 2 kilocycles, and to drop off at a rate considerably in excess of 6 decibels per octave at higher frequencies. The average spectrum slope of many typical machinery sounds seems to be about 12 decibels per octave in the region above 2 kilocycles. Therefore, most machinery sounds are too faint at supersonic frequencies to be detected with supersonic listening gear.[c] The only common exception to this observation is the type of sound produced by machines used to move fluids, such as pumps, compressors, or diesel exhausts. Such sibilant or explosive sounds are often fairly intense at frequencies well above 2 kilocycles and occasionally have appreciable supersonic content. Exhaust noise resembles automobile backfire. When diesel exhausts vent below the surface, the intensity of the underwater sound received from this source is very much greater than when they vent in

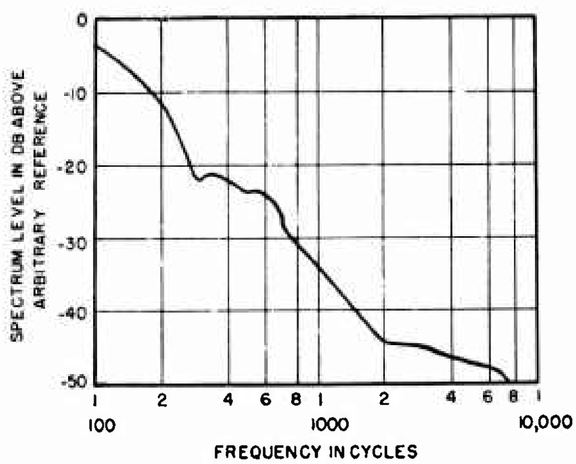


FIGURE 6. Spectrum of an S-class submarine at anchor, charging batteries with diesels.

air. Properly designed silencers may minimize the total output, however.

Finally, there is a miscellaneous group of noise sources which are occasionally important. This group includes (1) the slapping of waves against the hull, which is a function of sea state, course, and speed; (2) the bow wave, which is a function of hull contour and speed; (3) "body rumble," which arises from the rattling of plates and other loose structures, and

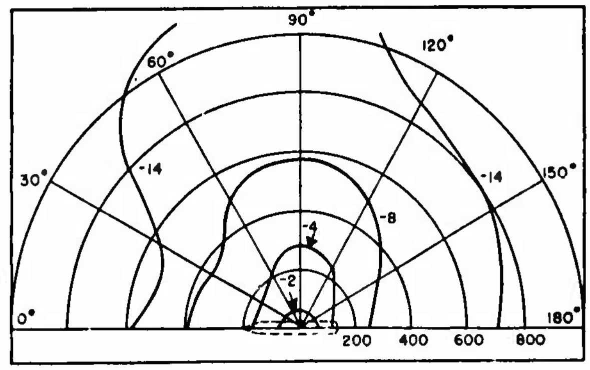


FIGURE 7. Equal-intensity contours for a 14-knot destroyer, measured in the band between 200 and 400 cycles. The contours give pressure levels in decibels relative to the maximum level observed beneath the ship; distances are measured in feet in the horizontal plane.

from the fact that a ship's hull is a drum which is subjected to various kinds of percussion; and (4) crew activities, such as voices, footsteps, hammering, and the ringing of gongs. This miscellaneous group generally produces a very small part of the total output, but it is important in a number of ways. Noise from crew activities, for example, may endanger a submarine taking evasive action; similarly, body rumble of the listening vessel may hamper sonic detection of targets.

All these various sources are generally distributed over the surface and through the interior of the hull. The underwater sound field near a ship generally changes with the bearing of a measuring hydrophone as well as with its distance from the ship. Figure 7 shows contours of equal intensity for sound in the 200- to 400-cycle band produced by a destroyer under way. Acoustic shadows cast by the hull and the wake would be expected to have precisely this effect on the contours of radiated sounds; such shadows should be better defined at high frequencies than at low. This effect is shown in Figure 8, which gives the contours measured near a moving freighter in the 2,500- to 5,000-cycle band. The contours in Figure 8

[c] In supersonic listening gear the frequency of the received sounds is reduced by heterodyning, so that the signal presented to the operator lies in the audio-frequency region.

are centered astern and presumably represent radiation predominantly from the propellers; contours of engine room noise tend to be centered amidships, as in Figure 7. In general, the intensity of propeller noise measured at

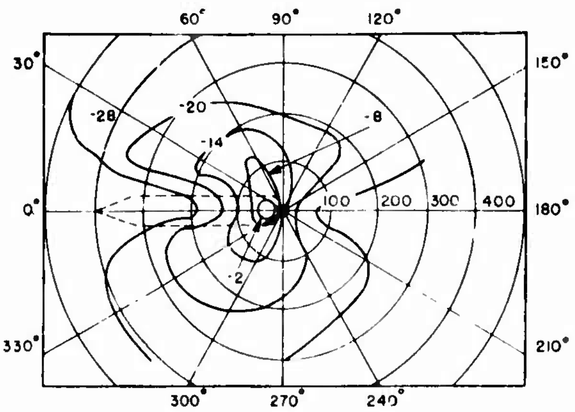


FIGURE 8. Equal-intensity contours for an 8-knot freighter, measured in the band between 2.5 and 5 kilocycles. The contours give pressure levels in decibels relative to the maximum level observed beneath the ship; distances are measured in feet in the horizontal plane.

a given distance from the propeller is greater for a receding than for an approaching ship. This dependence of intensity and composition on aspect (see Figure 9), together with changes in the intensity of cavitation due to acceleration, often enables sound operators to tell when a vessel changes course. Similarly, a change in the number of propeller beats indicates a change of speed.

Thus the acoustic outputs of ships are mixtures of sounds, and these mixtures can be put together in a large number of distinctive and easily recognizable combinations. Sound operators with experience in a particular harbor have been known to baffle the uninitiated through their uncanny ability to describe the superstructures of approaching ships by mere listening. This capacity of listening gear helps to classify unseen or poorly visible sources, since the resemblance in output between two sources of the same type usually outweighs individual differences, and it gives such gear obvious practical advantages over detectors which cannot identify sources.

The chief properties of the signal which help to detect and identify it are intensity, frequency composition, and time pattern (changes in the signal with time). Intensity and composition are shown by the spectrum of the signal, but spectra usually represent time averages and do not reveal time patterns present in the signal. The time patterns in ship signals are essentially variations of intensity, frequency, and phase. The ear may be very sensitive to variations in any of these, and to auditory motion in general. It was, for example, indicated in Section 2.2.2 that the ear is most sensitive to intensity changes occurring at a regular rate of about three per second, and that it is less sensitive to slower or more rapid changes. The several kinds of auditory motion may occur singly or in a great variety of highly individual combinations. Figure 56 in Chapter 4, for example, depicts the time pattern in cavitation noise produced by a submarine at periscope depth. This pattern is due to propeller modulation and has a period of about 2 beats per second. Figure 45 in Chapter 4 shows amplitude modulation in a tonal component in the spectrum of a tanker. These traces repre-

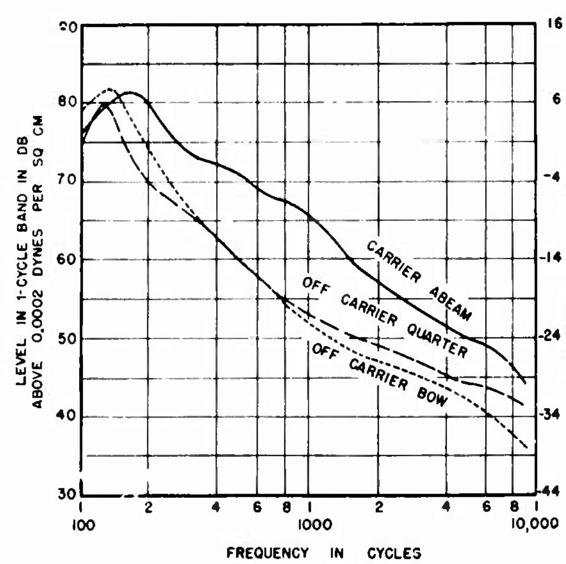


FIGURE 9. Spectra of an aircraft carrier measured at various aspects. Beam measurements were made 120 yards from the source; others were made at a distance of 240 yards.

sent variations in rms level in a band of indicated width, as measured by a power level recorder with a high writing speed. It is evident that the mean value of the rms level may differ by several decibels from the maximum and minimum values of rms level. As shown in Sections 4.2.3 and 5.1, detectability may often be determined by maximum and minimum rather than average values of the rms level.

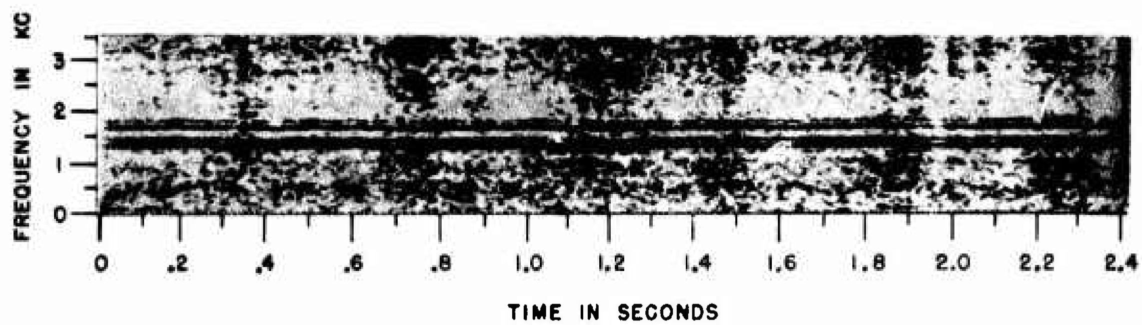


FIGURE 10. Time-frequency-intensity analysis for sonic noise from a 5-knot cargo vessel.

Figure 10 shows a record of the sounds produced by a cargo vessel moving at 5 knots. This record was obtained by sampling the various frequencies very rapidly with a 45-cycle band-pass filter. The midfrequency of this filter increased almost linearly with time; hence a vertical section of the record represents the intensities of nearly simultaneous samples of narrow-band noise taken over a large range of frequencies.

This record has three indicating scales and is one form of spectrum which shows time pattern. The scale of the vertical axis represents the midfrequency of the filter; the horizontal, the passage of time; and the darkness of the trace, relative intensity. The relative darkness is adjusted, in these traces, to compensate for an intensity loss of 6 decibels per octave. The successive vertical bands are propeller modulations occurring at a rate of about 3 per second; the nearly continuous horizontal striations represent tonal components at 1,350, 1,475, 1,725, and 1,850 cycles per second. The absence of propeller modulation from these tones indicates that they were probably produced by the propulsion machinery.

Figure 11 is a record of the noise produced by the same vessel moving at 10 knots; here the machinery tones are more nearly submerged by cavitation noise. Comparison of the two sets of traces shows that all the frequencies in cavitation noise are about equally affected by propeller modulation (see, however, Figure 12) and that the modulation rate increases with speed.

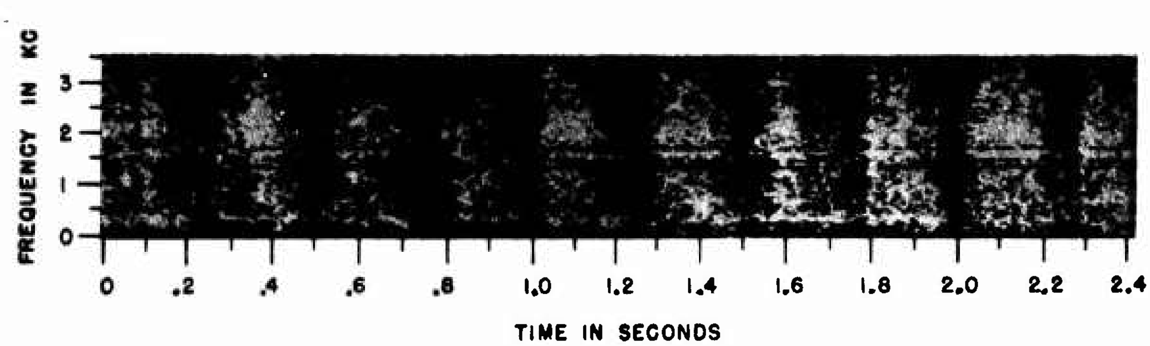


FIGURE 11. Time-frequency-intensity analysis for sonic noise from a 10-knot cargo vessel.

3.1.2 Far from Source

Sound may be changed in many ways during transmission through the sea.[d] The following

[d] Many of the effects that occur are described in STR Division 6, Volume 8, Chapters 1-10.

summary of these changes is adequate for present purposes. The signal grows progressively weaker with increasing distance from the source because of spreading and attenuation. The transmission loss is greater for the higher frequencies; in other words, selective attenuation occurs. In any actual case, the received signal is likely to show erratic fluctuations produced by skip-distance effects, interference patterns, and other factors. Figure 13, for example, is a record of changes in the level of sounds received in a stationary hydrophone and produced by an approaching tanker.

Stable interference patterns are considered improbable at distances greater than 2,000 yards and at frequencies higher than 2 kilocycles, since fixed phase relations between sounds traveling along two separate paths are not likely to persist beyond these limits. Such effects may nevertheless be important in the

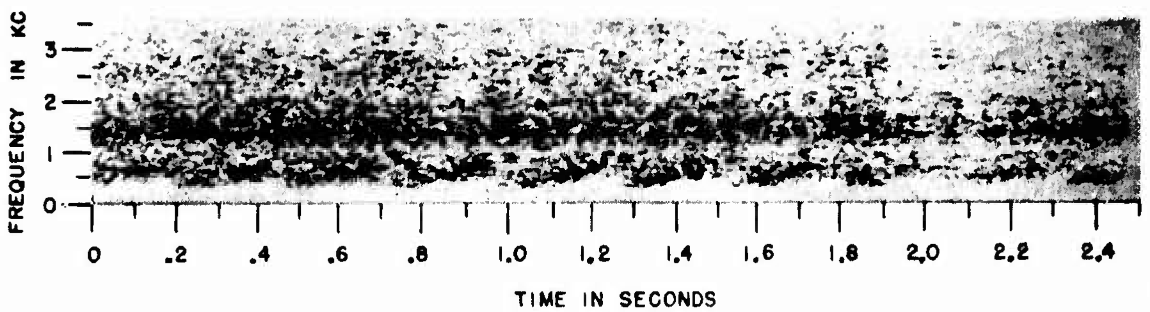


FIGURE 12. Time-frequency-intensity analysis showing frequency modulation of the propeller sounds from a submerged submarine. The modulation occurs in the region of 0.5 kilocycle and coincides with the period of the propeller blades, rather than that of the shaft. The swish was a distinctly audible component of the underwater sound.

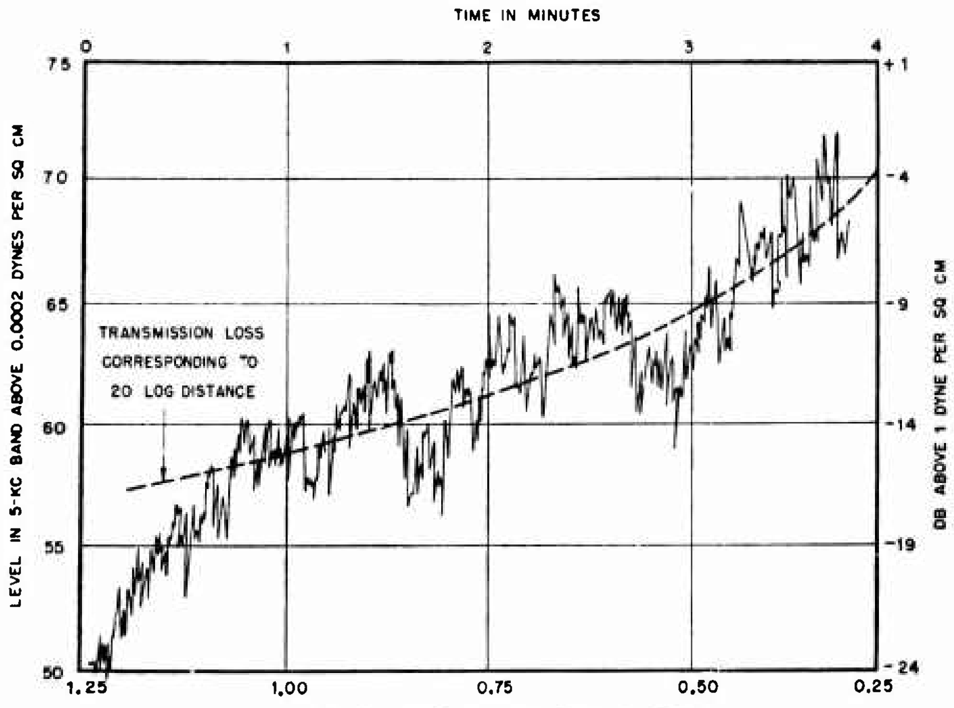


FIGURE 13. Power level record of the signal from an approaching tanker.

case of a submarine taking evasive action. Unstable interference effects, produced by continual shifts in phase between sounds arriving by slightly different paths, seem to be responsible for the rapid variability of transmission loss which is observed at all distances and all frequencies. These rapid fluctuations in transmission may have adverse effects on the operation of sound gear. The time scale of these fluctuations is generally different from that of propeller modulation and does not seem to hinder the counting of propeller turns.

A ship is an extended source; it is consequently possible to distinguish bow from stern, even at fairly long distances, by means of the change in quality of the sounds which are heard when the axis of a directional hydrophone is swept across the target bearing.

3.2 INTERFERING BACKGROUND

Many unwanted sounds enter the ears of sound operators. Since such sounds may mask faint signals, they set a range limit beyond which targets cannot be heard. The major sources of interference encountered in the listening gear of surface vessels and submarines are described in the present section.

Each of the sounds radiated by the listening vessel is a potential source of interference; hence the hydrophone should be shielded from them as carefully as possible. Since listening gear is tactically more effective when it can be used under way, it is difficult to attempt reduction of interference by suspending the hydrophone at a distance from the hull. Suspending cables tend to be unwieldy, and arrangements for steering the hydrophone to different bearings are not simple.

In most installations used up to the present time, the hydrophone is mounted just outside the hull; its cables and suspensions are led through the hull and are shock-mounted to reduce vibration noise. The effectiveness of the hydrophone varies with its position on the hull. It should, for example, be well forward from the propellers and the engine room; it should be located where it will receive the least interference from the auxiliaries and the bow wave; and it should be placed where it will be shielded from the impact of bubbles carried by the slipstream, but not where the hull will shield it from signals. On rapidly moving vessels, it has been found useful to enclose the hydrophones in a streamline dome to reduce the noise arising from the flow of water around the hydrophone.

The interfering sounds described so far are associated with the listening vessel and are generally termed *self-noise*. Self-noise increases with speed; it may be higher, for a given speed, when the vessel is in shallow water, as a result of reception of bottom-reflected sounds.

Interference may also arise from electrical noise, from airborne noise, and from ambient noise. *Electrical noise* refers to voltage variations produced by stray electromagnetic fields or by random thermal fluctuations in the hydrophone or its associated circuits. When these variations are amplified, they appear at the output as noise. *Airborne noise* does not originate in the listening system. It reaches the operator from his immediate environment and may even be produced, in exposed locations, by the effect of air streaming past the headphones.

Ambient noise consists of waterborne sounds which are not produced by the listening ship. It is a catch-all term applied to a large mixture of sounds. Some of the sources which contribute to this mixture are ships other than the target, waves and surf, rain, marine life, icebergs, and underwater volcanoes.

The level of ambient noise produced solely by motion of the sea surface (so-called *deep-sea ambient*) represents the lower limit of interfering background. Deep-sea ambient is essentially nondirectional; the sounds come with equal strength from all bearings. The intensity of deep-sea ambient noise rises with increase of wind force and sea state.

3.2.1 Sources

Self-noise seems to be the dominant source of interference in sonic listening gear.[c] Am-

[c] Sonic listening gear is designed to receive frequencies below 10 kilocycles; supersonic, to receive frequencies above 10 kilocycles.

bient noise is rarely a problem, except in heavy seas or when the listening vessel secures all its engines and auxiliaries. On the other hand, the level of self-noise received in the supersonic listening gear of slowly moving vessels is comparable to that of deep-sea ambient noise, and, in high seas, supersonic ambient noise may exceed self-noise. It is usually possible to keep electrical and airborne interference to a fairly low level in listening gear on ships and submarines.

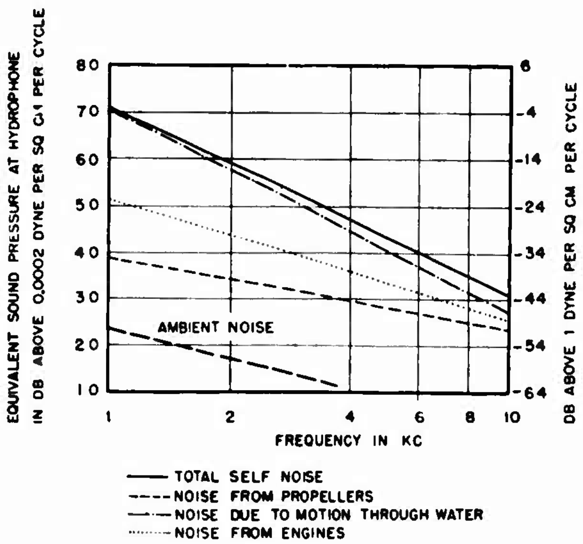


FIGURE 14. Components of sonic self-noise received in the listening gear of a small surface craft at 9 knots. Relative receiver bearing was 000, sea state 1.

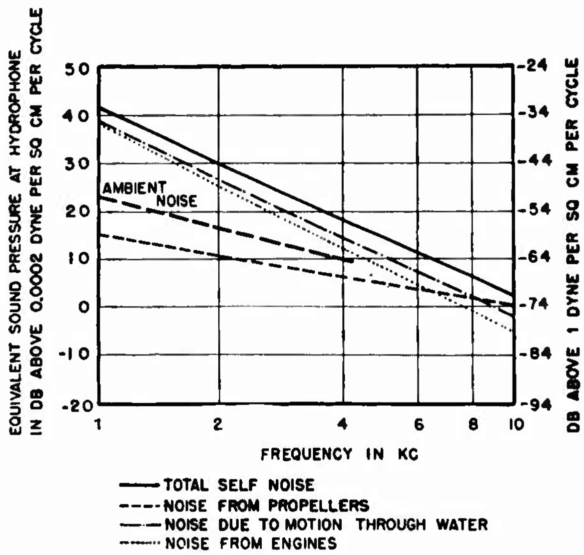


FIGURE 15. Components of sonic self-noise received in the listening gear of a small surface craft at 4 knots. Relative receiver bearing was 000, sea state 1.

The slopes of self-noise spectra are usually steeper than those of deep-sea ambient noise. This difference in slope tends to make self-noise relatively less prominent in supersonic gear than in sonic. Self-noise levels measured under way are highest when the hydrophone is pointed toward the propellers; for other orientations, the levels are lower and approximately independent of orientation.

Figures 14 and 15 give some observed spectra showing the relative magnitudes of noise components which were present in sonic listening gear installed on a small surface vessel. It is interesting to note the change in importance of propeller noise at the two speeds shown and the tendency for total noise to approach ambient as the frequency increases. These spectra are plotted to show rms levels at various frequencies in a hypothetical plane wave incident along the acoustic axis of the hydrophone; such a sound field, if it existed, would produce the same effect at the output as was actually observed. This is a convenient procedure because it represents all the noise sources on a common basis and in such a way as to simplify comparison with a signal spectrum, but it is not intended to imply that such a noise field actually exists at the hydrophone. There is a distinction, therefore, between the sound-in-the-water (that is, at the hydrophone) and the sound-at-the-ear; the sound meant will be specified in the remainder of this volume in all cases where the intended sense is not clear from the context. It is worth emphasizing that Figures 14 and 15 are presented for the sake of illustration only and should not be considered typical of all sonic gear.

Few background sounds have really distinctive character. Ability to detect and classify elusive differences in background character is the mark of a competent operator, and it pays dividends. Noting changes in the sounds received from his own ship's propellers and auxiliaries is an important function of the sound operator on submarines, since such changes often mean increased detectability by enemy listening gear. This requires good memory and good discrimination for intensity, pitch, quality, and time pattern. Obviously, no operator can detect the slight change in background which is produced by a faint signal unless he is thoroughly familiar with the background; conversely, an operator who consciously searches the background for features which resemble ordinary ship signals is less likely to report nonexistent signals.

Some backgrounds contain more popping and crackling sounds than others. Such impactive sounds are present in all backgrounds to some extent but are almost universal in the ambient noise observed in tropical or subtropical shallow-water areas over rock or coral bottoms. These sounds have been compared to the sound of frying fat; they are produced by colonies of snapping shrimp. Shrimp crackle is an important source of noise because it is intense, incessant, and widely distributed. The spectrum of shrimp noise shows a broad peak between 10 and 30 kilocycles; the level diminishes in the sonic range, and merges with deep-sea ambient in the neighborhood of 1 to 2 kilocycles. Ability to recognize shrimp crackle is important because it severely limits the ranges which can be obtained with supersonic gear; thus, shrimp noise may provide ideal acoustic camouflage for a slowly moving submarine.

3.3 PROPERTIES OF LISTENING GEAR

Studies of the performance of listening gear are described in Chapters 4 and 5. The following summary of the properties and uses of listening gear provides a basis for evaluating the results of these listening studies.

The chief function of listening gear is the detection and tracking of targets. Listening gear which can detect faint signals, in other words, which can detect targets at long range, is generally preferable to listening gear which cannot. On the average, longer detection ranges mean a greater number of contacts per cruise; they also mean that more time and space are available for maneuvering. Within the range of validity of the inverse square law, the statement that "one listening system can detect a target at twice the range that another can" is equivalent to the statement "the first system can detect a signal which is 6 decibels weaker"; in other words, 10 log $(d/2d)^2 = -6$ decibels, where d is the range obtainable with the second system.

It is difficult to open or close the range effectively or to compute torpedo courses unless the target bearing is known with adequate precision. The bearing error which may be tolerated with sound gear depends on the tactical situation and on the information obtainable from other types of equipment, such as radar or periscope.

It has already been pointed out that an estimate of target bearing may be made by crossing the target with the axis of a directional hydrophone; and that, at relatively short range, the bearings of bow and stern can often be determined. The qualitative change in the sounds which are heard during this transit permits the operator to judge the angular limits within which the target bearing seems to lie.[f] The midpoint of the arc defined in this way is taken to represent the target bearing. The difference in degrees between the true bearing and the estimated bearing is the bearing error. The magnitude of the probable bearing error depends on the hydrophone directivity and on various other factors (see Section 5.1).

The directivity of a hydrophone is determined by the width of the cone of nonaxial sound rays for which the hydrophone response is of the same order of magnitude as its response to an axial ray (see Section 5.1). When a small, distant source is moved to various points on the surface of a sphere circumscribed about a stationary hydrophone and the hydrophone response corresponding to various angular positions of the source is represented on either a spherical or a polar plot, the resulting figure is called a directivity pattern. A similar figure is obtained when the source is stationary and the hydrophone is rotated about its own center. For most practical purposes it is sufficient to show a two-dimensional directivity

[f] In other words, time pattern can be introduced in the received signal by crossing the bearing with a directional hydrophone. Thus, there are three types of time patterns: those due to the source, those resulting from transmission, and those produced by the operator.

pattern representing the response in the horizontal plane, that is, the plane described by the acoustic axis of the hydrophone when the latter is rotated about its mechanical (vertical) axis.

acoustic insulator (see Figure 2 in Chapter 5).

Two aspects of these diagrams are worth noting. First, there are usually several sectors where the response is high; these are known as *lobes*. By proper design, the magnitudes of

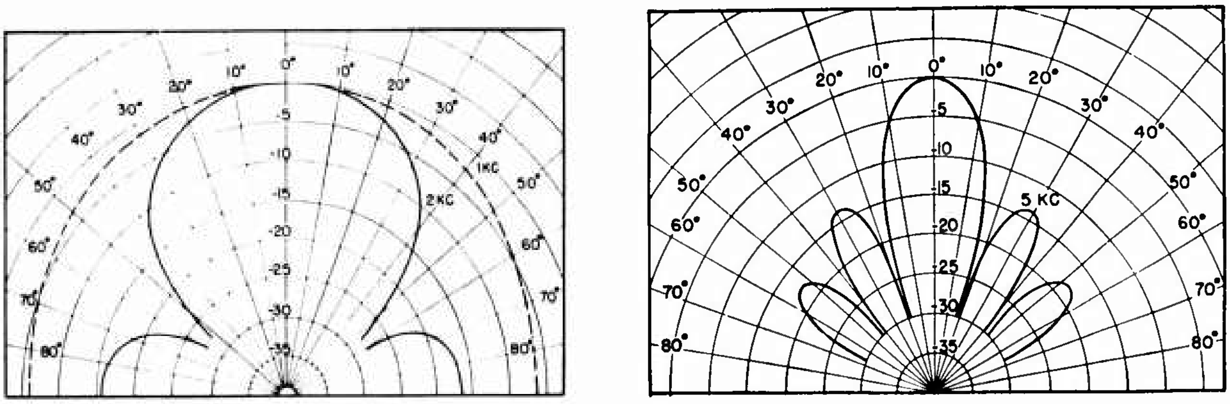


FIGURE 16. Theoretical directivity patterns for a 3-foot line hydrophone at 1, 2, and 5 kilocycles. The response is in decibels relative to axial response.

Figures 16 and 17 show the horizontal directivity patterns of two standard hydrophones. The sensitive elements of these hydrophones are assembled differently. In one case

the secondary lobes can usually be made considerably smaller than that of the primary or main lobe. Thus, operators listening to a target at maximum range are unlikely to hear it on

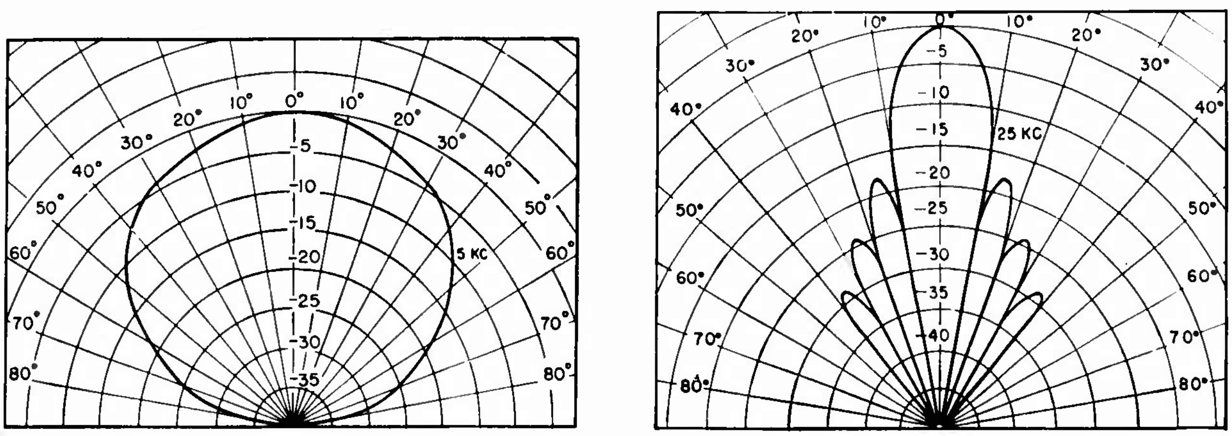


FIGURE 17. Theoretical directivity patterns for a 15-inch piston-type hydrophone at 5 and 25 kilocycles. The response is in decibels relative to axial response.

(piston type), the receiving elements are arranged on the surface of a circular disk; in the other (line hydrophone), along the surface of a narrow rectangle. In the case of the piston, the mechanical axis coincides with a diameter; in the case of the line hydrophone, with the minor axis of the rectangle. These figures show the front response of the hydrophones in decibels below the maximum response; the rear response is diminished by use of a baffle, or

a side lobe at one bearing, and again on the main lobe on another bearing;[k] the reason for this is that the background picked up on the main lobe will mask the relatively faint signal which can be picked up by orienting a side lobe toward the target. When the signal is well above background, it will be heard on the side

[k] The bearing scale is usually calibrated in terms of angle between axis of the main lobe and some reference bearing, such as dead ahead or compass bearing.

lobes and may even give a reciprocal bearing when the limit of insulation afforded by the baffle is reached. When a loud target is picked up at several bearings, the operator can decide which is correct by comparing the relative loudness of the signals received at different bearings (the loudness difference is usually sufficiently great to make such a decision easy) or by reducing the gain so that signals received on the side lobes approach inaudibility.

Secondly, it will be observed that the angular width of the main lobe varies inversely as the frequency of the sound for which the pattern is drawn. Thus, the available bearing accuracy will be greater at higher signal frequencies. If the perceived signal is distributed through a wide band of frequencies, the low-frequency components will show relatively little change of intensity when the hydrophone crosses the bearing, but there will be a comparatively large change in the intensity of the high-frequency components. In other words, the quality of the sounds heard will change when the target bearing is crossed; the change of quality heard in this case is due to the difference between signal and signal-plus-background. Thus the operator can achieve greater bearing accuracy, when wide-band signals are heard, by focusing attention on the higher frequencies or by using high-pass filters. If a wide-band signal is weak, however, the use of filters may reduce it to inaudibility (see Section 4.1.6). When the signal is strong, it may be dominated by its low-frequency components. Since these show a smaller intensity change when the target is crossed, and, therefore, give less bearing accuracy, it may be desirable to reduce the loudness of the low-frequency components of strong signals by dropping the amplifier gain. If the signal is sufficiently strong, reduced amplification also serves to eliminate false bearings picked up on the side lobes and to sharpen bearing accuracy by increasing the contrast between the faint background which is heard off-bearing and the louder sound which is received on-bearing.

When a directional hydrophone is placed in a nondirectional sound field, the hydrophone responds well to sounds incident on its axis and discriminates against nonaxial sounds. This reduction of background afforded by hydrophone directivity improves performance by making it possible to detect weaker signals. The extent of discrimination against nondirectional noise increases with frequency. Figure 18 gives the discrimination, against nondirec-

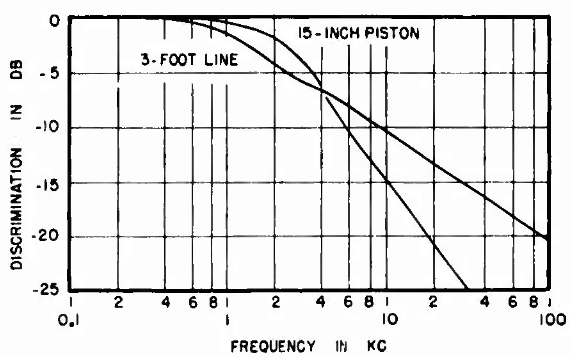


FIGURE 18. Discrimination of hydrophones against nondirectional noise.

tional noise at various frequencies, which is obtained with the hydrophones whose directivity patterns are shown in Figures 16 and 17. The discrimination is plotted in decibels below the response which would be obtained at each frequency if an equally sensitive nondirectional hydrophone were placed in the same sound field. This number of decibels represents essentially the ratio between the solid angle defined by the prominent parts of the hydrophone's response pattern and the total solid angle about a point, 4π steradians.

The total discrimination of a hydrophone against nondirectional sound is the sum of that shown in Figure 18 and that provided by the baffle. The latter suppresses approximately half of the response pattern (the rear response) for most hydrophones at most frequencies; less suppression is provided at low frequencies, because of diffraction, and near the resonance frequency of the baffle. On the average, the baffle adds about 3 decibels of discrimination.

Hydrophone directivity improves with increasing frequency and with increasing hydrophone dimensions. Since there is an upper limit to the size of practical hydrophone installations, supersonic listening gear usually gives better bearing accuracy and better discrimination against nondirectional background than can be obtained with sonic listening gear. Large sonic hydrophones have been constructed by connecting a group of small units in series. Figure 24 in Chapter 4, for example, indicates the discrimination against nondirectional background which is obtained from an assembly of 48 hydrophones arranged in the form of a ring 8 feet in diameter. The acoustic axis of such an assembly can be oriented by introducing various amounts of electrical delay[1] in the lines connecting the individual units. Thus, if the output of units closer to the source are delayed an appropriate amount, they are brought into phase with the output of units farther from the source. In this condition, the total output of the assembly is a maximum; in other words, the hydrophone is on-bearing. Hence, target bearing may be determined from the amount of delay needed in the various lines.

Figures 19 and 20 illustrate the mechanism of hydrophone directivity for an idealized case. Here, XW represents a progressive plane wave, of wavelength λ, incident on a hydrophone face YZ. The wave moves in the direction OM, M being the midpoint of the face, and d its length; thus, the angle of incidence is θ. When θ is zero, all the parts of the surface YZ move in phase, and the hydrophone output is maximal. As θ increases, a situation like that shown in the figure occurs; in other words, elements such

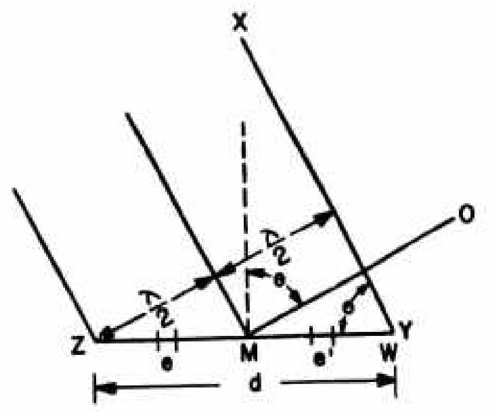


FIGURE 19. Mechanism of hydrophone directivity.

as e and e', selected from the two halves of the hydrophone and symmetrically placed with respect to its midpoint, are out of phase by 180 degrees. Under these circumstances, the resultant output represents the net effect of elements which cannot be paired in this way. When the effect of all the elements on one half face completely cancels the effect of the elements on the other half face, the output falls to zero. As shown in the figure, this occurs when $\sin\theta$ equals λ/d, that is, when the

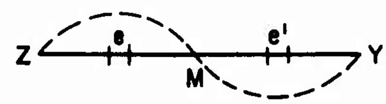


FIGURE 20. Phase of disturbance shown in Figure 19.

hydrophone face intercepts all the parts of the disturbance corresponding to a single wavelength. As θ increases beyond the value shown in the figure, the symmetry existing between the two halves of the hydrophone diminishes; hence cancellation is incomplete and the output is increased. The output then reaches another relative maximum because of increased reinforcement among the various elements and then falls to a second minimum when θ assumes such a value that the hydrophone intercepts exactly two wavelengths, that is, when $\sin\theta = 2\lambda/d$. In other words, there are a group of response lobes, the minima occurring whenever an integral number of wavelengths is intercepted by the hydrophone, that is, minima occur for all values of θ, in the expression $\theta = \text{arc sin } (n\lambda/d)$, which correspond to integral values of n.

The width of the main lobe is simply the angular distance between the first minimum on one side of the axis and the corresponding first minimum on the other side. Since the first minimum falls at an angle θ such that $\sin\theta = \lambda/d$, the width of the main lobe is $2 \text{ arc sin } (\lambda/d)$. When λ is considerably smaller than d, the angle θ is equal to its sine, and thus the width of the main lobe is $2\lambda/d$ radians. If c/f is substituted for λ, where c is the velocity and f the frequency of the sound incident on the hydrophone, the lobe width equals $2c/fd$ radians. In other words, the lobe width, and hence the response to nonaxial sounds, decreases when either the dimensions of the hydrophone or the incident frequency is increased. It is evident that when λ exceeds d, the phase differences across the hydrophone are negligibly small, and the hydrophone loses almost all directionality.

The magnitude of the response along the axes of the side lobes is smaller than that along the axis of the main lobe ($\theta = 0$ degrees) because some cancellation always occurs between the two halves of the hydrophone face when θ exceeds zero. Thus, response patterns for a given hydrophone show an increasing number of side lobes for higher frequencies, but all the lobes are narrower for the higher frequencies, and the response on the side lobes is usually very small compared with that obtained on the main lobe. A directional hydrophone will therefore pick up a smaller fraction of the high-frequency energy in a nondirectional sound field than of the low-frequency energy; in other words, it has a greater degree of discrimination against high-frequency sounds.

3.4 TYPES OF MASKING

The preceding discussion has been devoted to the separate characteristics of signal and noise spectra as received by listening gear. Before discussing the observed masking data, it may be helpful to survey the problem broadly and to examine various types of masking situations that may arise. First, however, the concept of primaudibility, which is basic in the study of masking, is defined.

3.4.1 Definition of Primaudibility

In order to distinguish between *threshold-limited listening* (audibility determined by sound level relative to the absolute audibility threshold) and *masking-limited listening* (audibility determined by level of signal relative to a background which itself lies above the audibility threshold), it has been suggested that the coined word *primaudible* (from *prime* plus *audible*) be used to describe the faintest signal which is audible under masking-limited conditions. This general term should be qualified to denote the percent detection probability which is intended (such as "50 per cent primaudible"); where it is not qualified in the present discussion, a detection probability of 50 per cent is understood. The term *audible* is thus reserved for threshold-limited listening.

Since detectors other than the ear may be used with sonar equipment, it has been proposed that the word *priceptible* (from *prime* plus *perceptible*) be employed as a general term to describe the faintest signal which can be perceived by the detector under masking-limited conditions, and that the term be qualified as above with respect to detection probability. Since these definitions make for brevity and precision, their use has been adopted in this volume.

3.4.2 Masking by Single Tone

In this case, the masking curves of Figure 2 in Chapter 2 apply. When the signal is a pure tone whose frequency approximates that of the masking tone or its subjective harmonics, the cue to recognition is essentially a fluctuation in loudness resulting from beats between the two tones. The minimum perceptible intensity increment determines the ratio of signal to noise at primaudibility.

For a pure tone of other signal frequencies, the limiting factor is competitive stimulation of the basilar membrane by the masking tone (remote masking). The amount of this stimulation can be determined only by empirical methods at the present time, but a number of observed regularities have been listed in Section 2.1.2. The effects of frequency and amplitude modulation have not been studied for the present case.

When the signal has a continuous spectrum, the criterion for audibility is presumably that the signal intensity in at least one of the ear's critical bands be equal to or greater than the intensity of the single tone which would be primaudible at that frequency. No detailed laboratory evidence is available, however, to test this expectation.

3.4.3 Masking by Continuous Background

If the background has a continuous spectrum, the limiting factor is competitive stimulation of the basilar membrane, and the critical band criterion applies. Recognition should then be possible when the signal energy in at least one critical band is equal to the background noise energy in the same band. When the isolated signal components are widely different in pitch and character, the critical band criterion applies to each component individually. In other words, the various signal components do not necessarily cooperate to improve signal audibility; in fact, it is possible to hear one component and miss the other, or others (see Sections 4.1.3 and 4.1.5). When the signal components lie within a narrow frequency interval, the group may or may not be more audible than the individual components.

When signal and background spectra have similar compositions within a frequency interval extending over many adjacent critical bands, the cue is usually change in loudness, and the condition which determines primaudibility is the size of the intensity limen (least perceptible intensity increment).

3.1.4 Masking by Group of Tones

The case in which the masking background, as well as the signal, consists of a group of tones has not been studied. It will be obvious, however, that all the masking situations already discussed above are limiting cases of this last situation and may be approximated by increasing or decreasing the number and the frequency separation of the tones in the signal and background. It is possible, therefore, that useful estimates of the masking of a group of tones by another group may be obtained by interpolating among observations made on the preceding simpler types.

Chapter 4

LABORATORY MEASUREMENTS ON MASKING OF TARGET SOUNDS

Listening gear provides a means for achieving certain practical objectives. The extent to which current gear may be expected to meet the objectives set for it and the likeliest methods for getting better results depend in part on the ability of the human ear to detect target sounds. While the basic information on the structure and performance of the ear, described in Chapters 1 and 2, helps to answer many questions, the extent of the ear's ability to detect underwater sounds generated by surface vessels, submarines, torpedoes, and other sources can best be decided by experimental test and analysis. This program has made considerable progress but is, in many ways, still incomplete. Chapters 4 and 5 set forth its major findings and indicate briefly some of the things which remain to be done.

The purpose of the listening tests described in these chapters is to determine how well the ear of an average, competent operator can perform certain assigned tasks. The general nature of these tasks and the broad features of the hearing process have been outlined in the preceding discussion.

Two general types of test have been conducted: (1) laboratory tests, in which the field situation is simulated as faithfully as seems necessary and practicable, and (2) field tests, in which the listener is stationed aboard ship and the listening situation closely approximates the one encountered in practice. The laboratory measurements, made under controlled conditions, are described in this chapter, while the field measurements are discussed in the following chapter. Both types of test provide useful information, and both types have characteristic merits and defects. Field tests, for example, can be made extremely realistic, but they are time-consuming; also they require that personnel, marine facilities, and favorable weather be available simultaneously. Furthermore, field tests yield rather uncertain conclusions unless (1) the factors which determine performance are accurately measured and controlled, or (2) a statistically significant series of tests is made. These requirements are difficult to meet in the case of field tests. They are readily met in laboratory tests, but such tests are inevitably artificial. Hence considerable care is necessary in the design and interpretation of laboratory tests if practically useful information is to be obtained from them.

Listening tests are intended to provide definite, quantitative information. In other words, such tests are essentially measurements; like other measurements, they must be properly planned and performed if they are to be helpful rather than misleading. Thus, the sounds used in the tests should be typical of those met in practice. The observers should have normal hearing, and, if possible, the same or comparable observers should be used throughout a test series. The observers should understand the situation and know exactly what they are expected to do. When the test situation is new to the observers, their performance usually improves during the time needed to learn the ropes. It is desirable to make provision for such preliminary adjustment in order to approximate the performance to be expected from trained personnel. Even after such preliminary adjustment, the test observers may not do as well at the assigned task as would experienced field personnel. It seems reasonable to expect, however, that any advantage in favor of personnel accustomed to a particular routine would be offset by various factors, such as fatigue, boredom, and distractions, which may affect performance under service conditions but which do not usually enter the test situation. Finally, precautions must be taken to eliminate guessing and practice effects, such as memorizing the test items or anticipating the order in which they are presented.

The preceding requirements apply to all test programs; they also apply to programs designed to train personnel in the effective use of sound gear. Other points of procedure will emerge in the following discussion of observations. These observations are described in semi-chronological order, and the discussion is interspersed occasionally with general material not contained in the original reports. This method of presentation has been adopted because of dissimilarities in the techniques and materials of observation which were used at different laboratories. Although most of these dissimilarities are of minor significance, it would burden the discussion to keep all the distinctions to the fore at all times. Specific details of apparatus and test administration and analysis are generally omitted.

4.1 BRITISH TESTS

This group of tests[1-4] was a pioneering study undertaken to examine the factors affecting aural detection of sonic submarine sounds masked by deep-sea ambient noise. Among the factors studied were the choice of observers, method of presenting the sounds used, gain setting, filters, and discrimination of directional gear against nondirectional background. The apparatus used is shown schematically in Figure 1. Recordings of the signal and the background, suitably amplified, were fed to a mixing panel and, after further amplification, the mixture of sounds was presented to the observers through headphones. In a given series of tests the setting of the background amplifier, as well as that of the mixture amplifier, was maintained constant, and the setting of the signal amplifier varied in steps of 2 decibels. The observers were seated in a quiet room. A selected mixture of sounds was presented to them for 15 seconds, and, in the silent interval before the next mixture was presented, they recorded whether the signal was "heard" or "not heard." This procedure determined the signal-to-noise ratio needed for detection. The required measurements of signal and noise levels were made with the calibrated measuring amplifier shown at the right in Figure 1, and the results obtained are presented in Figures 6 through 23. The information accompanying these figures describes the source of each sound and its characteristics as heard in the absence of the masking background. The chief feature of interest in these figures is the relation existing between the spectrum of the masking background and the spectrum of the signal which can barely be detected. In general, they indi-

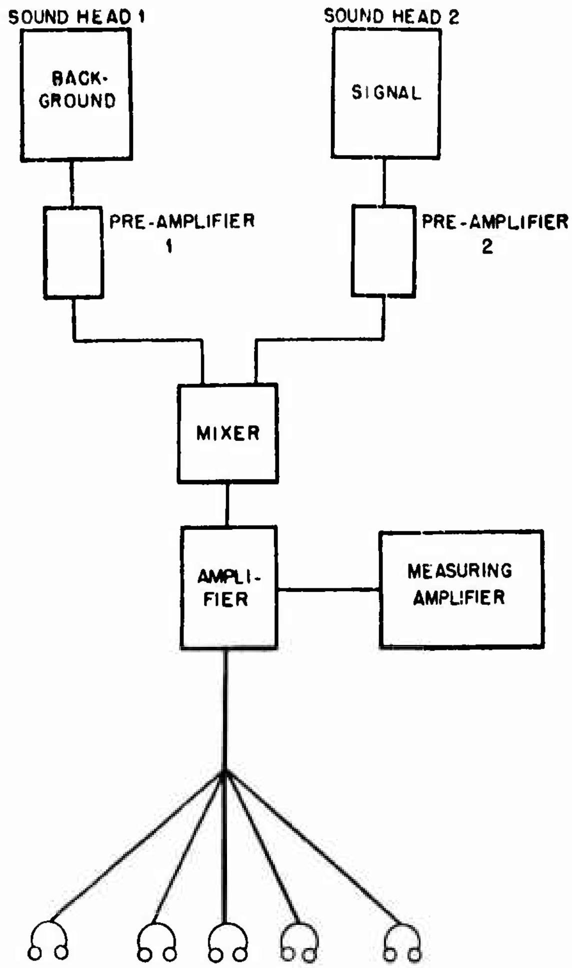


FIGURE 1. Schematic diagram of listening test apparatus.

cate that recognition of a primaudible wide-band signal tends to occur at the frequency where the signal spectrum level is highest relative to background. Other significant aspects of the observations are discussed in the following sections. It should be noted, however, that a number of the terms and concepts used in the discussion of this group of tests were not explicitly stated prior to publication of the work described in Section 4.2.

4.1.1 Nature and Measurement of Sounds

The noises used were obtained from film recordings of submarine and water sounds. These recordings were made at a sound range and were obtained with a system whose overall response was quite flat from 60 to 8,000 cycles. The recording hydrophones were close to the submarine, so that all but the faintest submarine sounds were well above background noise present in the water. Representative sections of the films were selected, the ends of the sections spliced to form continuous loops, and the recorded sounds were reproduced by running the loops on one of the sound heads shown in Figure 1. The advantage of film over disk recordings, for work of this sort, is that film recordings are less susceptible to changes produced by wear. The use of recordings for listening tests minimizes the amount of time during which the target vessel is unavailable for other assignments.

Two different sets of recordings were used for these tests; those involved in Figures 6 through 15 were made at an earlier date than the ones described in Figures 16 through 23. The recordings of signal and background which were used in these listening tests were always obtained with the same recording system. Failure to observe this precaution may give quite unrealistic results when the system response is very different for signal and background recordings; for example, false peaks and hollows may be introduced into the spectra.

The use of an ambient-noise background for these tests corresponds to one tactical situation of interest, since antisubmarine vessels, with the aid of a cable-suspended hydrophone, sometimes conduct a search while adrift. Knowledge of the level at which a given submarine sound can be heard in the presence of water-noise background may then be interpreted in terms of the range within which a submarine can be detected by listening. This requires adjustment for the energy lost in transmission of the sound from submarine to searching vessel. Due allowance must therefore be made for such factors as selective attenuation and interference patterns produced by direct and surface reflected sound (see Section 3.1.2). Such range information is equally useful in prosubmarine and antisubmarine work. Roughly, a required increase of 6 decibels in signal level means the detection range is halved.

Several sources of error must be considered when the results of laboratory listening tests are used in detection-range calculations; these factors must also be considered, therefore, when the laboratory tests are planned. The signal used may not be typical of the class. It may be subjected to amplitude and frequency distortion if nonlinear systems are used in the recording or presentation of sounds; in particular, the mixture amplifier in Figure 1 may be troublesome in this respect, unless properly designed. Weak signals may contain large amounts of background noise. Thus, the high-frequency[a] output of the signal records used in these tests contained appreciable ambient noise, except for the cases shown in Figures 11, 12, 14, 16, and 17. This situation is artificial, since it may lead to "detection" through increased loudness of the recorded ambient noise.

One of the fundamental purposes which a program of listening tests can serve is to provide a method for predicting the audibility of signals on the basis of spectrum, time pattern, and other characteristics. The spectra of signal and background used in these tests were determined by measuring the fraction of the acoustic energy in the sounds which was transmitted by each of a set of octave band-pass filters. Spectrum level of the noise background is plotted in decibels below the overall energy, that is, the energy contained in all the frequencies between 50 and 10,000 cycles per second.

Octave filters are designed to transmit only the energy contained between the frequency limits f and $2f$, $2f$ and $4f$, $4f$ and $8f$, and so on To find the average energy per cycle, which is called the *spectrum level* when expressed in decibels, the ratio $I/\Delta f$ is formed, where I is the energy passed by the filter and Δf is the number of cycles between the upper and lower frequency limits of its pass band.

[a] For practical convenience in description, the following approximate sonic frequency intervals are distinguished: low frequencies are those below 2 kilocycles; middle frequencies include those between 2 and 5 kilocycles; and high frequencies extend from 5 to 10 kilocycles.

The relations between the arithmetic midfrequency F of an octave filter and the lower and upper cutoff frequencies f_1 and f_2 of that filter are shown in Figure 2. Since the pass

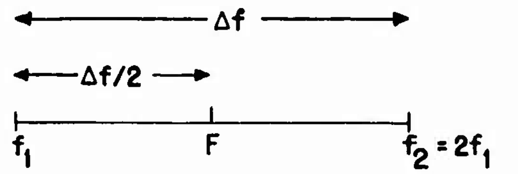


FIGURE 2. Characteristic of an octave filter.

$$\frac{\Delta f}{2} = \frac{f_2 - f_1}{2} = \frac{2f_1 - f_1}{2} = \frac{f_1}{2}, \tag{1}$$

hence

$$\Delta f = f_1. \tag{2}$$

If F is the arithmetic midfrequency, then

$$F = f_1 + \frac{\Delta f}{2} = \frac{3}{2} f_1, \tag{3}$$

and

$$f_1 = \tfrac{2}{3} F. \tag{4}$$

Then from (2) and (4),

$$\Delta f = \tfrac{2}{3} F; \tag{5}$$

in other words, the number of cycles in the pass band of an octave filter is equal to $\frac{2}{3}$ of its arithmetic mid-frequency.

band Δf of such a filter includes an increasingly greater number of cycles as the midfrequency F is raised, the use of octave filters shows progressively less detail for the higher frequencies in the spectrum. In fact, the pass bands of octave filters exceed critical band widths for all frequencies above 100 cycles (see Figure 17 in Chapter 2); octave filters are therefore poorly adapted to studies in which the ear's critical bands play a part.

When measurements made with octave filters are reduced to a "per cycle" basis, it is necessary to select a frequency, within the pass band of the filter, to which the computed value of the average energy per cycle is to be assigned. If the spectrum is flat, that is, if the energy per cycle is the same for all frequencies, it is sufficient to select the arithmetic midfrequency of the filter. In most noise spectra, however, the energy is concentrated in the lower frequencies; the common practice, therefore, is to assign the deduced average energy value to the geometric midfrequency $\sqrt{f_1 f_2}$ of the octave filter. As shown in Figure 3, this procedure is valid only for spectra which have a slope of -6 decibels per octave (when $I = k/f^2$, where k is a constant), but it is a fair approximation for other slopes as well. In obtaining the spectra shown in Figures 6 through 23, it has been assumed that (1) the filters used were "rectangular" (in other words, that they had sharp cutoffs), (2) the average energy per cycle should be assigned to the geometric midfrequency of the filter, and (3) the spectra contained no tones or other discontinuities. These

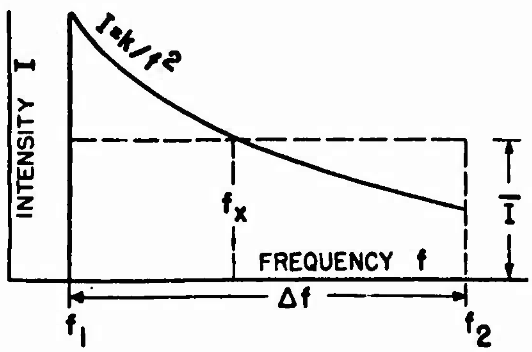


FIGURE 3. Effective midfrequency of an octave filter. The problem is to find the frequency f_X where the actual value of I equals the average value $\bar{I}$. Thus, if the area under the curve equals the area of the rectangle,

$$\bar{I}\Delta f = \frac{k}{f_X^2}\,\Delta f = \int_{f_1}^{f_2} \frac{k}{f^2}\,df = -\frac{k}{f}\Big]_{f_1}^{f_2}$$

$$= \frac{k}{f_1} - \frac{k}{f_2} = k\,\frac{f_2 - f_1}{f_1 f_2} = k\,\frac{\Delta f}{f_1 f_2}. \tag{1}$$

Therefore

$$\frac{k}{f_X^2}\,\Delta f = k\,\frac{\Delta f}{f_1 f_2}, \tag{2}$$

or

$$f_X = \sqrt{f_1 f_2} = \sqrt{2 f_1^2} = 1.414 f_1 \tag{3}$$

$$1.414(\tfrac{2}{3}F) = 0.94F.$$

assumptions are probably close to the truth for the case of the ambient background, but the last assumption is almost certainly untrue for all the signals, except cavitation noise (or "propeller thrash"). This is indicated by the recurrence of the words "hum" and "tone" in the descriptions of the signals. The energy content of these tones is unknown, but they must have been some 15 decibels above the adjacent spectrum level of the signal in order to escape masking (see Figure 17 in Chapter 2).

It will be noted, in Figures 6 through 23, that there is invariably a flat peak in the spectra in the frequency region where tones were detected. Such peaks would be expected if spectra contained tones superimposed on distributed sounds. Obviously, the effectiveness of

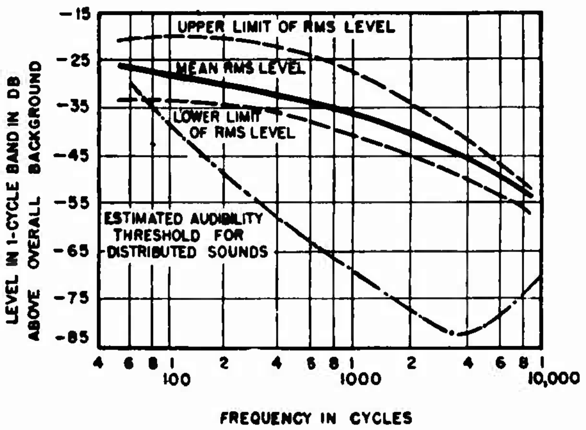


FIGURE 4. Fluctuations in water noise produced by occasional loud "plonks" due to waves lapping the side of the recording vessel. The mean level of water noise produces a loudness of 70 phons.

such a tonal peak in raising the total energy measured through a wide-band filter diminishes as the filter width increases.

Most of the energy measurements were made with a sluggish instrument, a copper oxide meter, and it should be noted that rhythmic or random peaks in sound level are not indicated very well by such a device. The peak factor (which refers, in the present discussion, to the difference in decibels between mean and peak values of rms level) for machinery sounds of the type shown, as estimated with the copper oxide meter, was found to be about 3 to 5 decibels. A similar peak factor applies to the background noise. Thus in Figure 4, the upper dashed curve represents occasional peaks in the background noise which were produced by waves lapping the side of the recording vessel. The spectrum levels corresponding to these "plonks" were obtained with a power level recorder. The continuous line in this figure shows the mean level of background noise during the intervals between wave slaps; the dashed lines show the upper and lower limits of deviations from this mean level. The dot-dash curve is discussed later.

The background spectra shown in Figures 6 through 15 represent the composition of the sound between wave slaps. The pitch of this sound changed somewhat during the course of the record. It is therefore represented in Figures 6 through 15 by two analyses, taken at the beginning and the end of the record, respectively, and the mean trend of these analyses is given by the continuous curve in Figure 4. Marked fluctuations in the levels of signal spectra are shown in Figures 11 and 15. It is interesting to note that the background spectrum given in Figures 16 through 23 does not exhibit such marked variability, presumably because the water noise shown was recorded with a hydrophone mounted on the bottom instead of near the surface.

The trend of the nondirectional background spectra shown in these figures is -6 decibels per octave. The spectrum level of the nondirectional background at 300 cycles is about 30 decibels below the overall level of the background. When applied to sonic spectra, the term *overall level* means the energy in the band from 0.1 to 10 kilocycles; this band of frequencies is termed the *standard reference band*. The frequency limits of the spectra shown in these figures very nearly coincide with those of the standard reference band. As indicated later, the limits of the standard reference band are imposed by the responsiveness of sonic listening devices and the sensitivity of the ear.

This relation between the overall level and the spectrum level at 300 cycles may be verified by using the method indicated in Figure 3, that is, by evaluating the energy in the standard reference band for various spectra and determining the fractional part of that energy in a 1-cycle band at 300 cycles. The relation is often useful in working with spectrum levels, and is a good approximation for all distributed spectra with constant slopes of -3 to -9 decibels per octave.

In contrast with the background spectra, the energy distribution of the machinery sounds shows no uniform dependence on frequency other than that it tends to peak at the low frequencies. In other words, machinery spectra tend to be individualized. It is interesting to compare the spectra of the sounds produced when the bow and stern planes are operated (Figures 6 and 8). Similarly, Figures 7 and 21 show the difference in output of ballast pumps on two submarines. The spectra of the sounds shown in Figures 13A and 13B are not so much typical of speech alone as they are of speech coming over a loudspeaker and transmitted through a submarine hull. Figures 14 and 15 indicate the character of the noise radiated at or below the cavitation threshold; Figures 16 and 17 show well developed cavitation, and Figure 18 illustrates an intermediate stage (see also Figures 4 and 5 in Chapter 3). Finally, Figures 19 and 20 show the increase in prominence and frequency of gear whine produced by increasing speed, and Figures 22 and 23 show a similar effect for pump noise.

The spectra reproduced in Figures 6 through 23 give the distribution of energy in the sounds used, as measured at the output of the mixture amplifier shown in Figure 1. This distribution is not identical with that occurring at the output of the headphones. In other words, the sound at the ear may be significantly different from the sound in the water.[b]

The factors which affect the performance of headphones, in addition to their own electrical and mechanical characteristics, are (1) the volume of enclosed air, (2) leakage around the cap, which is probably most significant for frequencies lower than 500 cycles, (3) the construction and resonances of the ear canal, (4) the yielding of the drum membrane and the walls of the canal, and (5) the directional character of the phones. In other words, the performance of headphones depends upon the impedance into which they work.

It should also be mentioned at this point that temperature changes encountered in practice may significantly alter the performance of headphones. Similarly, slight changes in the position of the phones with respect to the ears may produce large changes in the loudness of their low-frequency and high-frequency outputs.

[b] Failure to take this factor into account seriously weakens several of the conclusions in reference 28.

4.1.2 Estimated Audibility Threshold

Figures 6 through 23 show the relation existing between the spectra of signal and background when (1) the signal-to-noise ratio was sufficient to give detection in half the presentations, and (2) the loudness level of the received sound was observed to be approximately 70 phons, in other words, the intensity of the equally loud 1,000-cycle tone was 70 decibels above threshold. This is a comfortable loudness

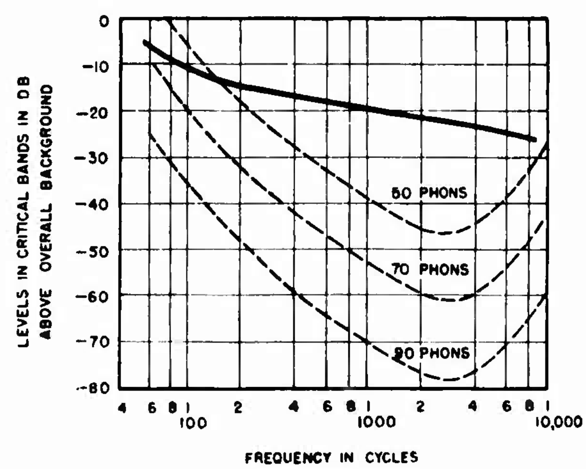


FIGURE 5. Estimated sensation levels of water noise, relative to the absolute threshold for tones. Solid line shows critical-band levels of average water noise relative to overall noise level. Dashed lines indicate positions of thresholds for pure tones when the presented water noise has the loudness levels shown on the curves.

for continuous listening. It is important to know the level above threshold for each frequency under the conditions of these tests. While this cannot be determined accurately for these British data, the analysis may be instructive.

The level of background relative to the absolute audibility threshold of an average observer taking these tests may be estimated by the methods mentioned in Section 2.3, in other words, by finding the fractional part of the total energy (in the mean water noise shown by the unbroken line in Figure 4) which is contained within the limits of each critical band stimulated by the presented sound, and summing the loudness contributions made by all the stimulated bands. The solid line in Figure 5 represents computed energy levels in critical bands at various frequencies and differs in shape from the 1-cycle spectrum of the background shown in Figure 4 because critical band width is a function of frequency (see Figure 17 in Chapter 2). The dashed curves in Figure 5 give the headphone threshold for pure tones (see Figure 1 in Chapter 2). To conform to the practice followed by the British, all curves in this section are plotted in decibels relative to the overall background. Thus, the relative position of the pure-tone threshold plotted on this basis will depend on the overall level; hence three threshold curves have been drawn to correspond to overall levels of 50, 70, and 90 phons.

The spectra in Figure 4 are represented on the basis of energy per cycle $I_{\sim}$. Hence, the energy I_{cb} available for stimulating a critical band of width Δf_{cb} is $I_{cb} = I_{\sim} \Delta f_{cb}$, or 10 log $I_{\sim}$ = 10 log I_{cb} − 10 log Δf_{cb}. As pointed out above, this is the relation connecting the average water-noise spectra shown in Figures 4 and 5. To a good approximation, the same expression describes the relation between the tonal threshold I_T at a frequency f, and the threshold intensity I_d of a distributed sound extending between the limits of the critical band centered at f; in other words, 10 log I_d = 10 log I_T − 10 log Δf_{cb}, where I_d expresses the energy per cycle when a distributed sound attains its threshold value. The quantity computed from the last expression may be termed the "threshold for distributed sounds." In Figures 6 through 23, which are plotted on a per cycle basis, this threshold for distributed sounds is shown for an estimated background loudness level of 70 phons.

It should be noted at this point that the spectra and the threshold shown in Figure 5 are to some degree inconsistent, since the composition of the water noise was measured at the input to the headphones whereas the illustrated tonal threshold refers to their output. In the absence of information on the properties of the headphones used in these tests, there is no quantitative basis for revision of Figure 5. If the response of the headphones used was not approximately flat over a large part of the sonic frequency range, the curves plotted in Figures 4 and 5 may have little relevance for these British tests.

In general, however, the illustrated relations are in qualitative accord with observations made in these and other tests. For example, the relation between the critical-band spectrum of water noise presented at 70 phons and the tonal threshold, shown in Figure 5, is in fair agreement with the data shown in Figure 76, which represents a similar situation, except that the measurements used in plotting both the threshold and the spectrum shown in Figure 76 were made at the input to the headphones. Thus, the estimates furnish a useful picture of the approximate conditions of test and emphasize one type of measurement which is required to supplement the straightforward masking studies.

Although audibility thresholds and headphone outputs have usually been ignored in listening tests, it is clear that these quantities may significantly affect the nature and interpretation of results, as well as the recommended design requirements. In the case at hand, it appears doubtful, on the basis of the estimated audibility threshold and the probable headphone response, whether much weight should be given to the spectra in the regions below 100 or above 7,000 cycles. Between these limits, however, the slope of the background spectrum seems well adapted to optimal listening; the level of the background noise above threshold does not change as much in this frequency range as would be the case for a flat background. Consequently, the background plotted in Figure 5 tends to produce the sensation of a "full" or "well-balanced" sound since it gives more nearly equal emphasis to all the frequencies than a flat background would give.[c] To obtain a completely "well-balanced" sound, the noise level in each critical band should produce the same subjective loudness. Thus, the decrease of the background relative to threshold

[c] Such a sound is often called a "white" noise, but this term is also used to describe a continuous spectrum with the same amplitude at all frequencies. Since it is not always clear from the context whether the expression is intended to mean "objectively" or "subjectively" white, spectra are usually described in the remainder of the discussion by means of their slopes, in decibels per octave.

shown for low frequencies in Figure 5 is probably desirable in order to give these frequencies equal loudness with the others.

The advantage of a well-balanced sound is presumably decreased operator fatigue. With such a sound it is equally easy for the operator to focus attention on each portion of the frequency band, and the high-frequency components are not annoyingly loud. The masking of sounds will not be directly affected by the slope of the background, unless this slope is so steep that remote masking occurs (noise in one critical band masks a sound in another). For most backgrounds, the noise frequencies masking the signal are only those falling within the same critical band. Thus the slope of the background, while it has some importance in operator fatigue and overall effectiveness, is not of crucial importance in the detection of target sounds.

The validity of this conclusion and of the critical-band criterion for masking of distributed sounds depends, of course, on the importance of remote masking. While little quantitative information on remote masking for distributed sounds is available, it seems probable, on the basis of Figure 6 in Chapter 2 and Figures 76 through 79, that this problem will not arise if the 1-cycle spectrum of the presented sound has a slope not exceeding some 25 decibels per octave, or in terms of Figures 6 through 23, if the slope of the presented sound is not significantly steeper than the slope of the audibility threshold for distributed sounds. Under such circumstances, the critical-band criterion for masking of tones by distributed sounds should apply without modification; this conclusion agrees with the results discussed in Section 4.2.5.

If the response of the headphones were flat and the loudness level were known to be accurately 70 phons, the signals plotted in Figures 6 through 23 would be audible only if they lay above the plotted thresholds. However, the overall background may have differed by as much as 5 decibels from the estimated 70 phons. In addition, the headphones may not have had a flat response, especially at the frequencies below 500 cycles, at which the signal tends to lie close to or below the plotted threshold curves. For this reason, no great reliance can be placed on these computed thresholds. They are included here primarily to indicate the general possibility of analyses along these lines, rather than to draw precise conclusions from the data.

4.1.3 Observed Recognition Differentials

The relation which existed between the signal and background spectra, when the signal was just recognizable, was determined by plotting a transition curve and adopting 50 per cent recognition as the criterion of detectability. The term "transition curve" refers to a plot of *percentage recognition* against signal-to-background ratio, where percentage recognition is defined as the fraction of trials at a given signal-to-background ratio for which the signal was heard. The signal and background energies are measured in the standard reference band, and their ratio is usually expressed in decibels. The transition curves obtained with each of the signals, under various test conditions, are shown in Figures 6 through 23. The signal-to-background ratio found in this way for a 50 per cent recognition probability is called the *recognition differential*, usually abbreviated as *RD*. The *RD* is commonly expressed as the decibel difference between overall signal and overall background. Thus the recognition differential gives the relation between overall signal and background levels at detectability. Since the filter analyses determine the spectrum levels for the various sounds in terms of decibels below the overall level of the sound, the relation between the different frequency components of the signal and background spectra is completely determined. Values of the observed recognition differentials are given in the figure legends.

From the preceding definition it is clear that the *RD* specifies a critical value of signal-to-noise ratio, and that the smaller the *RD*, the easier is detection. Since the detectable signal level is usually below background, the recognition differentials tend to be negative; any con-

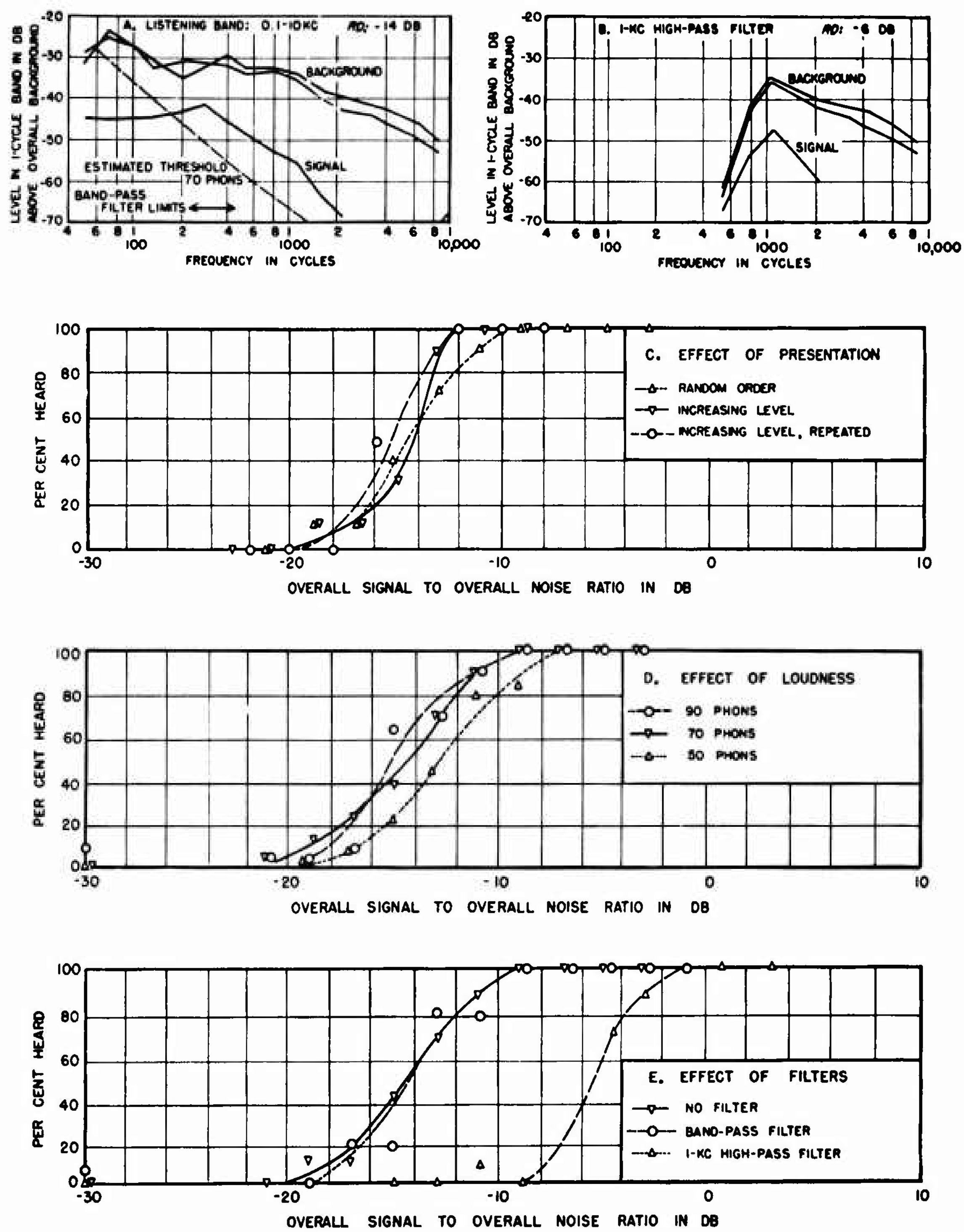


FIGURE 6. Audibility of sonic noise from aft planes (in power), masked by water noise. Character of signal: steady rapid knocking, with steady hum fluctuating slightly in pitch.

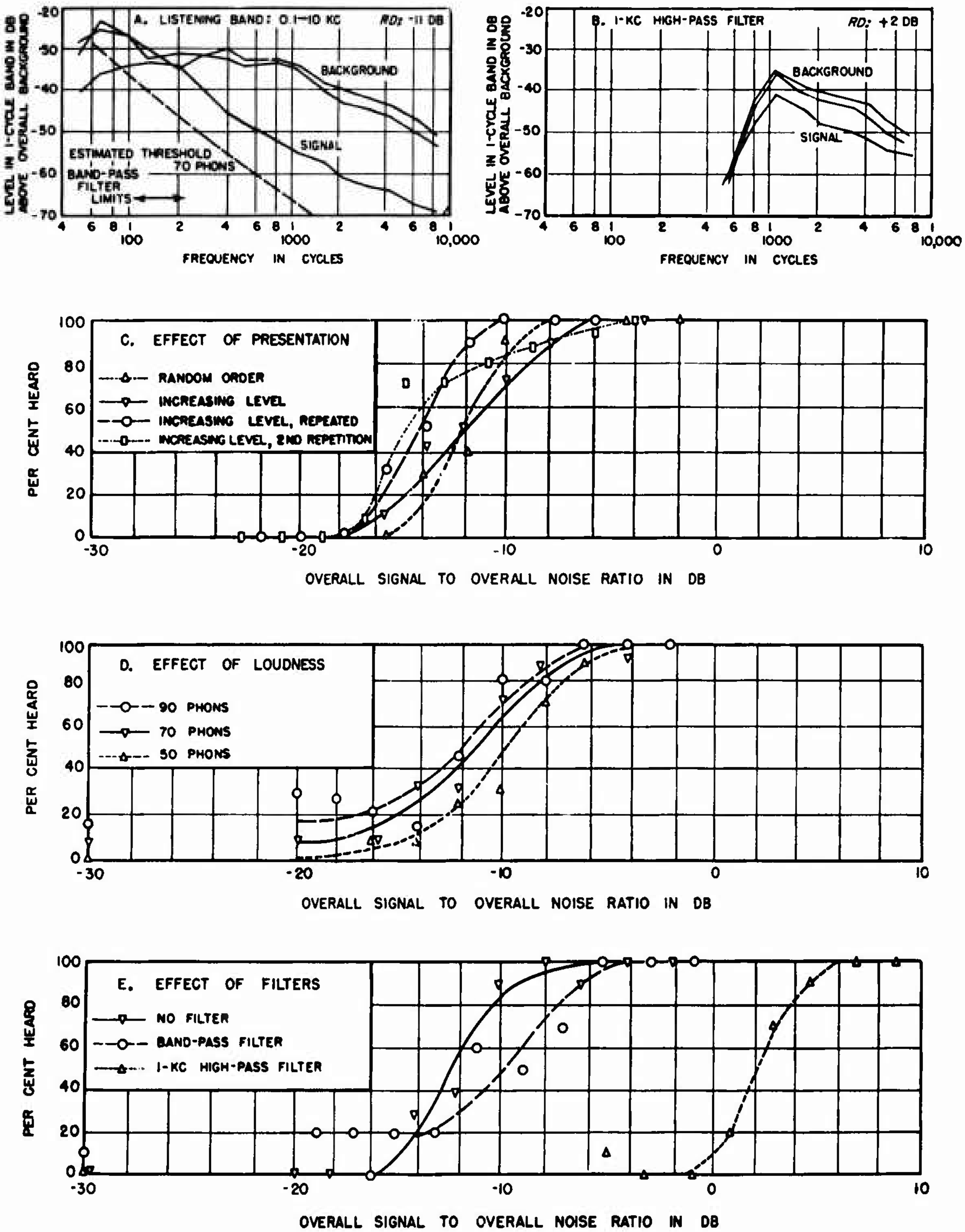


FIGURE 7. Audibility of sonic noise from ballast pump, masked by water noise. Character of signal: steady low-pitched hum and irregular crackling sound.

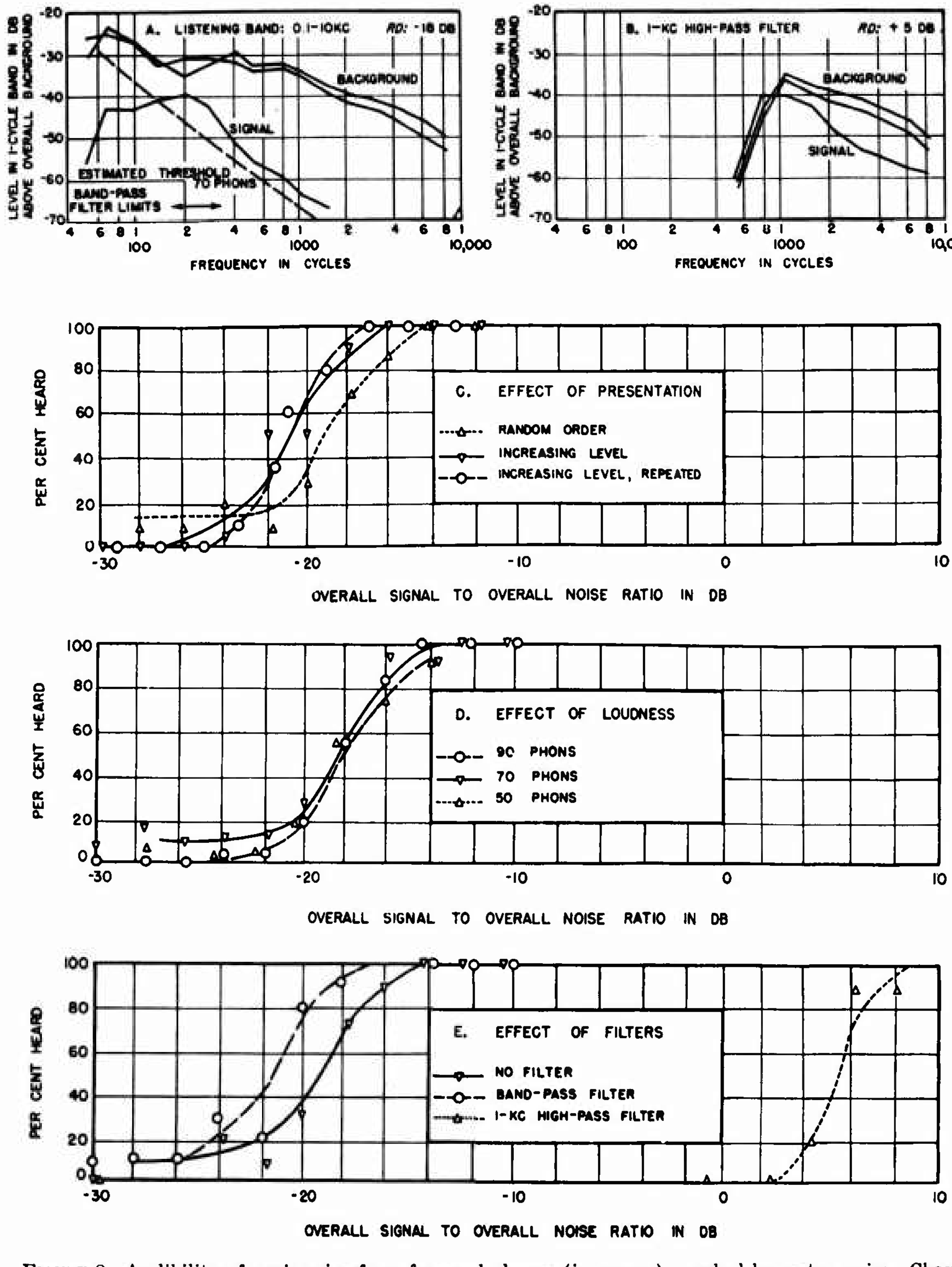


FIGURE 8. Audibility of sonic noise from forward planes (in power), masked by water noise. Character of signal: medium-pitched hum, with steady pitch and fluctuating intensity.

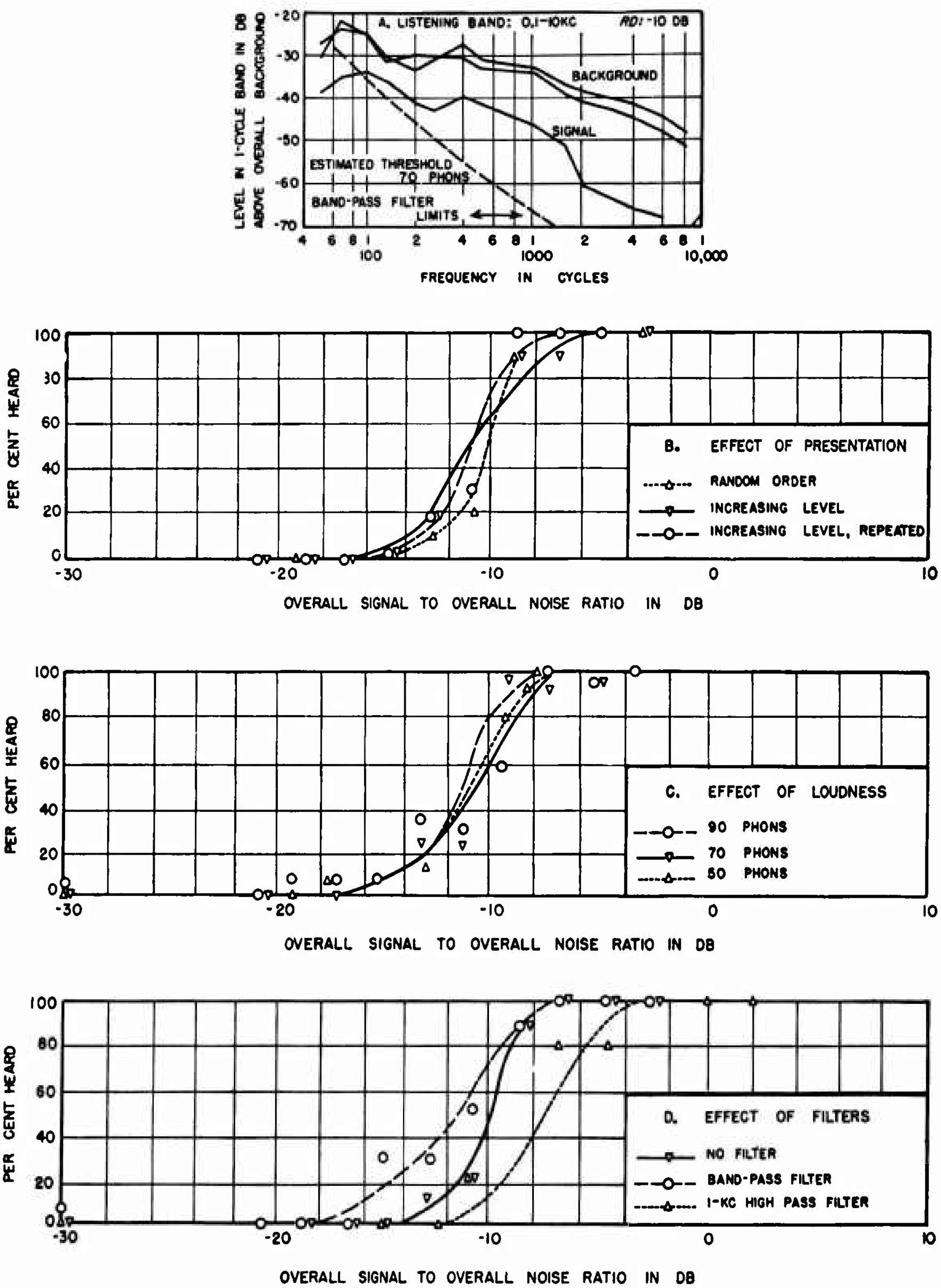


FIGURE 9. Audibility of sonic noise from starboard auxiliary circulator, masked by water noise. Character of signal: medium-pitched hum, with steady pitch and fluctuating intensity. A low-pitched hum was also audible.

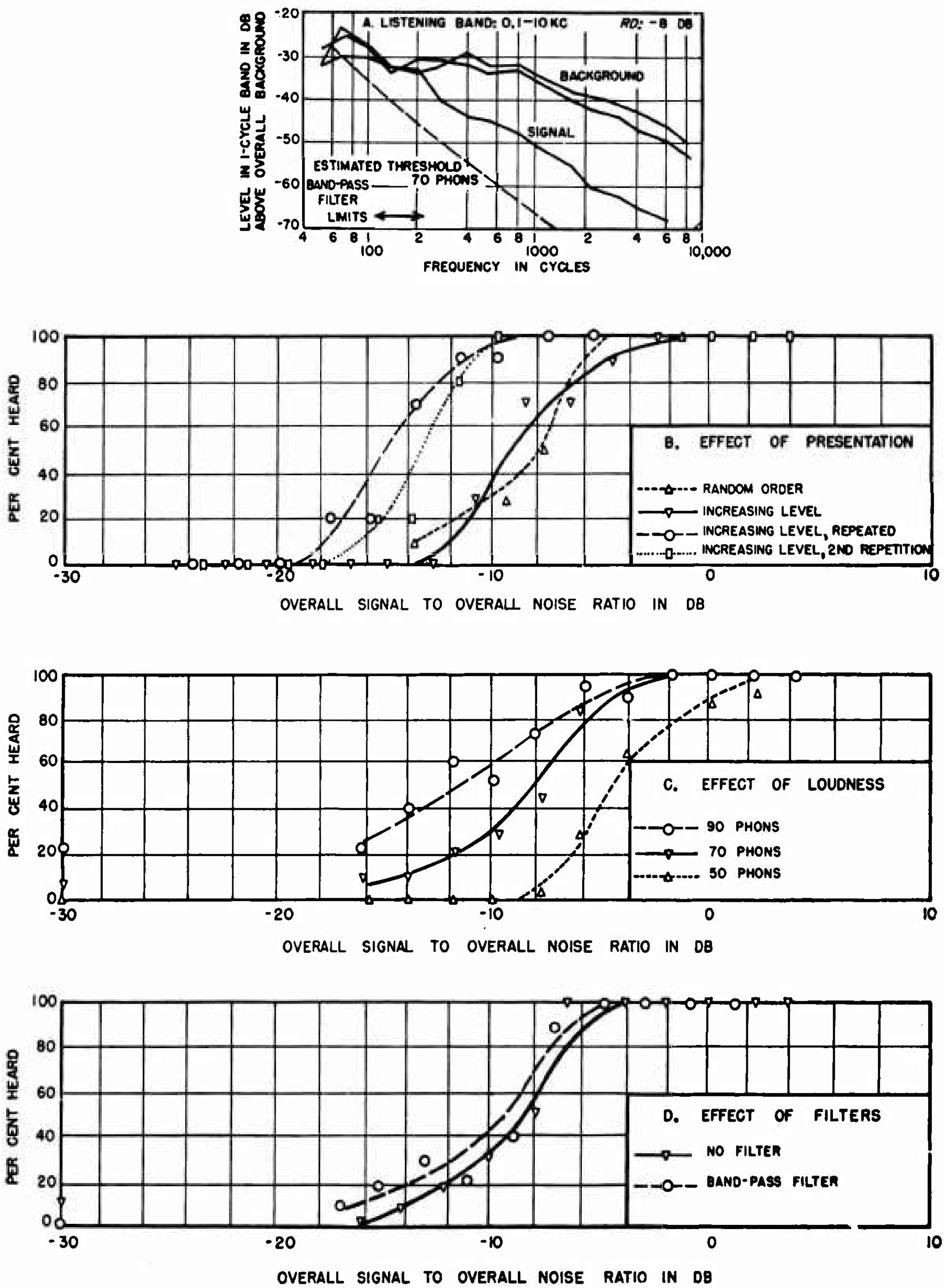


FIGURE 10. Audibility of sonic noise from port motor cooling fan, masked by water noise. Character of signal: low-pitched hum, with steady pitch and fluctuating intensity. Much water noise was also audible.

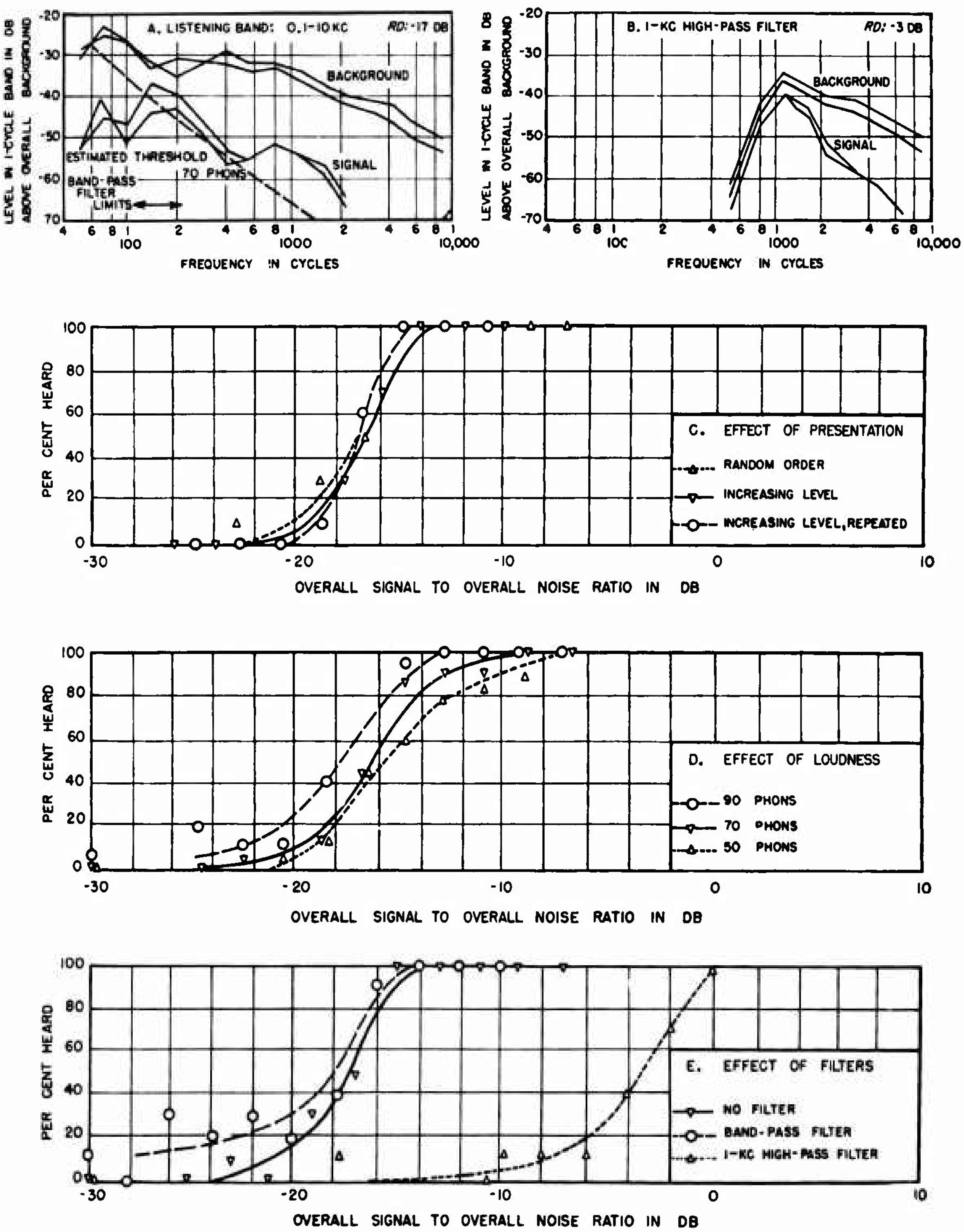


FIGURE 11. Audibility of sonic noise from lubricating oil separator, masked by water noise. Character of signal: medium-pitched hum, which rose in pitch and fell in intensity toward the end of the record. Also audible was a faint "ch ch ch" at 8 or 10 per second.

dition which interferes with detection yields a smaller negative number (a larger *RD*). Since the ear has a fairly wide dynamic range, the absolute levels of signal and background are define the frequency bands in which signal and background are measured. If the intensities of the sounds fluctuate, the time constant of the measuring instrument is a significant quantity;

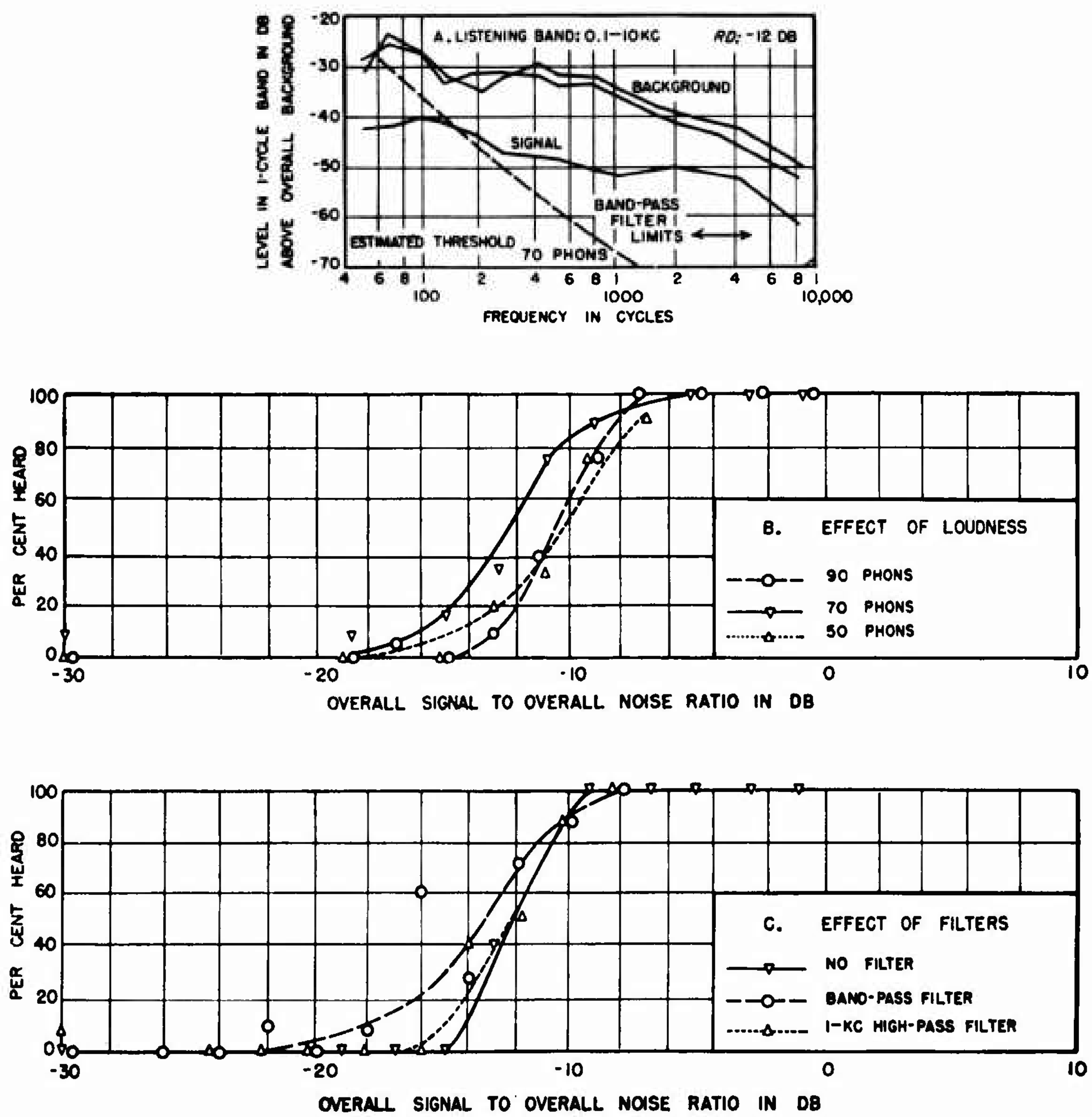


FIGURE 12. Audibility of sonic noise from lowering of aft periscope, masked by water noise. Character of signal: sharp "ch" twice during the record.

usually described in terms of subjective loudness only—for example, comfortably loud. Due to the fact that signal-to-background ratios at primaudibility are usually less than unity, the mixture of background and primaudible signal is only slightly louder, on the average, than the background alone.

For the sake of precision, various quantities should be known accurately. It is necessary to so also is the point in the fluctuation cycle, for example, peak, mean or low, which is used to determine the level. The limits of deviation among observers and among tests should probably be indicated. Similarly, the duration of each trial, the interval between trials, the total scope and duration of a test, and the subject's level of experience, acuity, and fatigue should be known.

Recognition differentials for the various submarine sounds shown here vary over a range of nearly 30 decibels. The extremes are illustrated in Figures 8B ($RD = +5$ db) and 16B

observers taking the same test or for a given observer repeating a test.

A much more reliable index of detectability is the difference in level between signal and

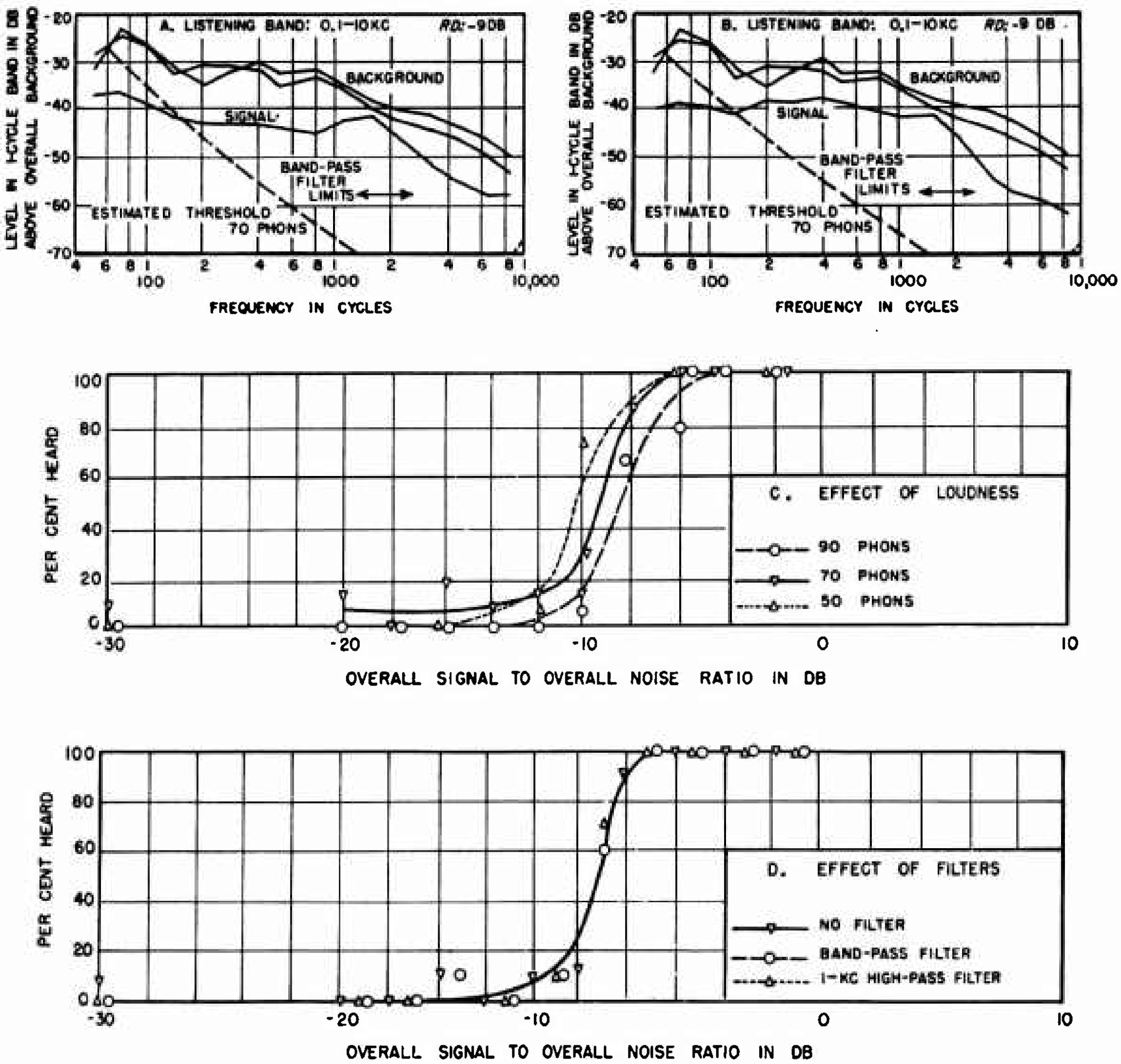


FIGURE 13. Audibility of speech transmitted by a submarine loudspeaker system, masked by water noise. Character of signal: words "six" to "seventeen". The signal was intelligible although much water noise was audible. Spectrum A applies to "six," spectrum B to "ten."

($RD = -24$ db). Even for a given signal under two different conditions of test (Figures 8A and 8B—see Section 4.1.6 for a description of the tests), the RD was observed to change from -18 decibels in one case to $+5$ decibels in the other. This high degree of variability is not due to erratic behavior of the ear, since the scatter among the recognition differentials was rarely more than 1 to 2 decibels for different

background spectra at the frequency where these spectra most closely approach each other. Thus, it will be observed that the total range of variation in this quantity is 10 decibels. The limits of variation are illustrated in Figures 7 and 9, where the difference in level at the frequency of closest approach is 0 decibel, and in Figure 18, where the spectra are spaced about 10 decibels apart at the point of closest ap-

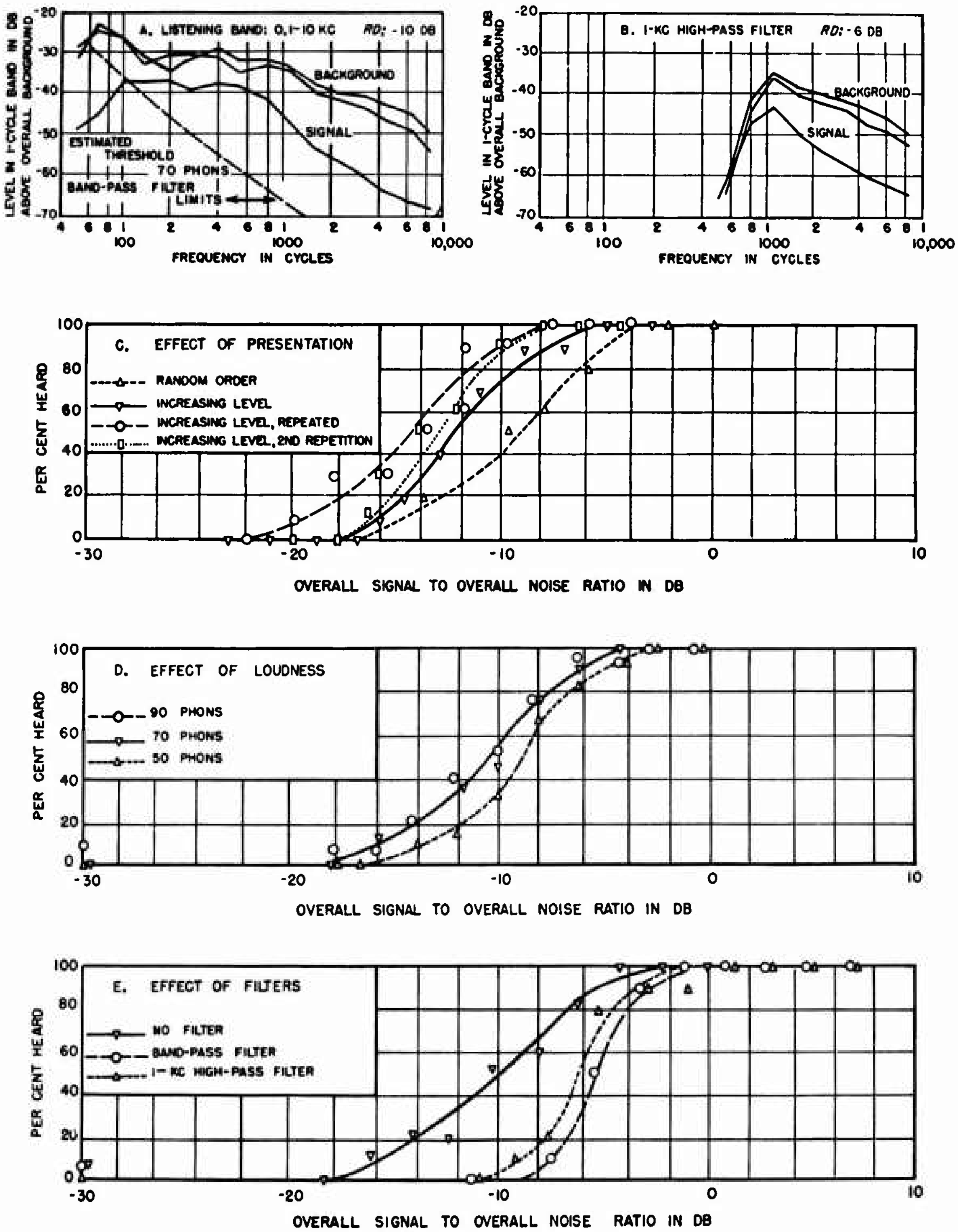


FIGURE 14. Audibility of sonic noise from port propeller (124 rpm), masked by water noise. This submarine proceeded at approximately 2.5 knots when both propellers were operated at 124 rpm. Character of signal: "ch ch ch" at 6 per second, and "swoosh" at 2 per second. The two modulation periods correspond to the blade and shaft rates, respectively, of a 3-bladed propeller. Also audible was a medium-pitched hum with steady pitch and variable intensity; a low-pitched hum was just audible.

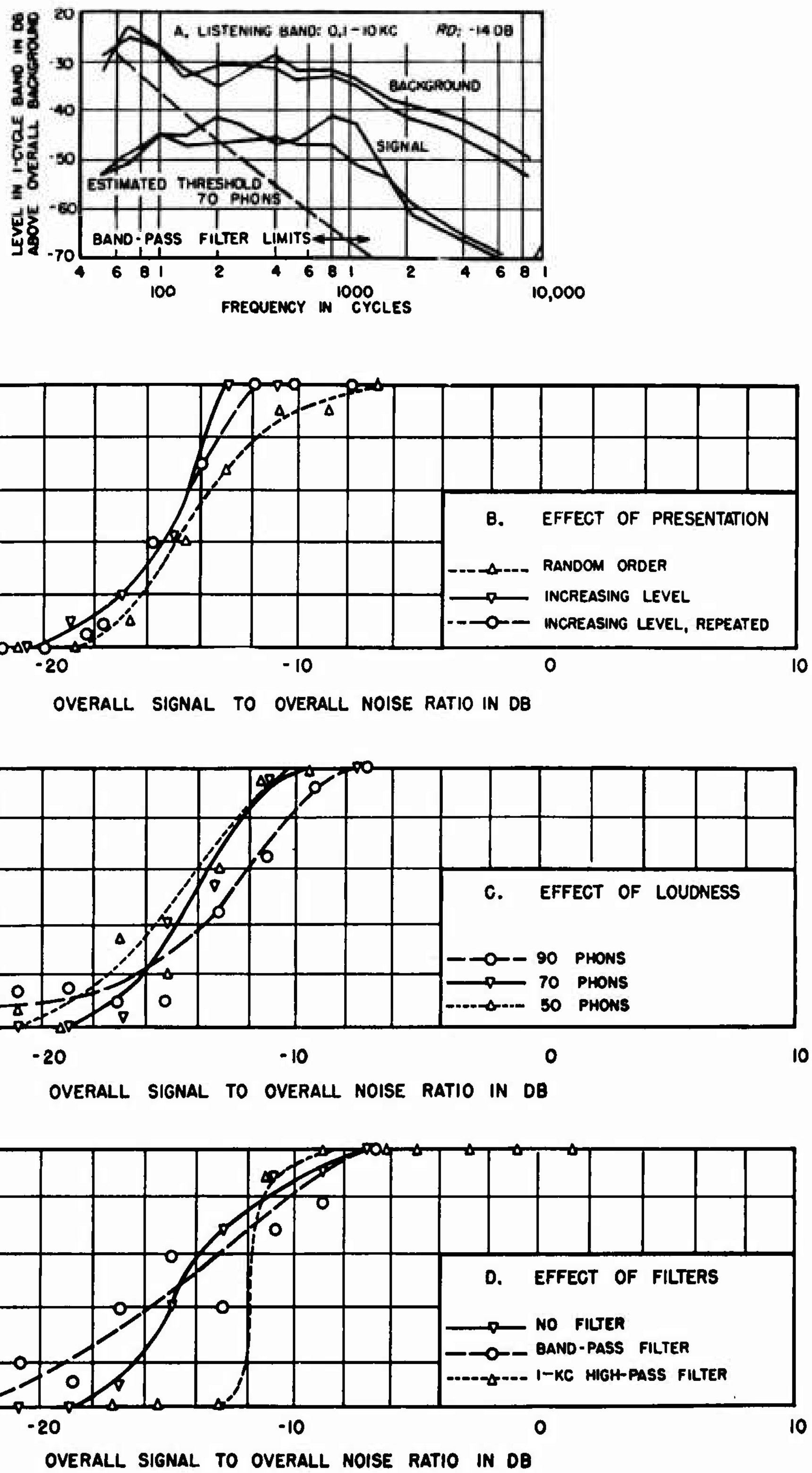


FIGURE 15. Audibility of sonic noise from main motors in series (120 rpm), masked by water noise. Character of signal: high-pitched hum with steady pitch and variable intensity; a low-pitched hum was just audible, but the signal contained no noticeable propeller noise.

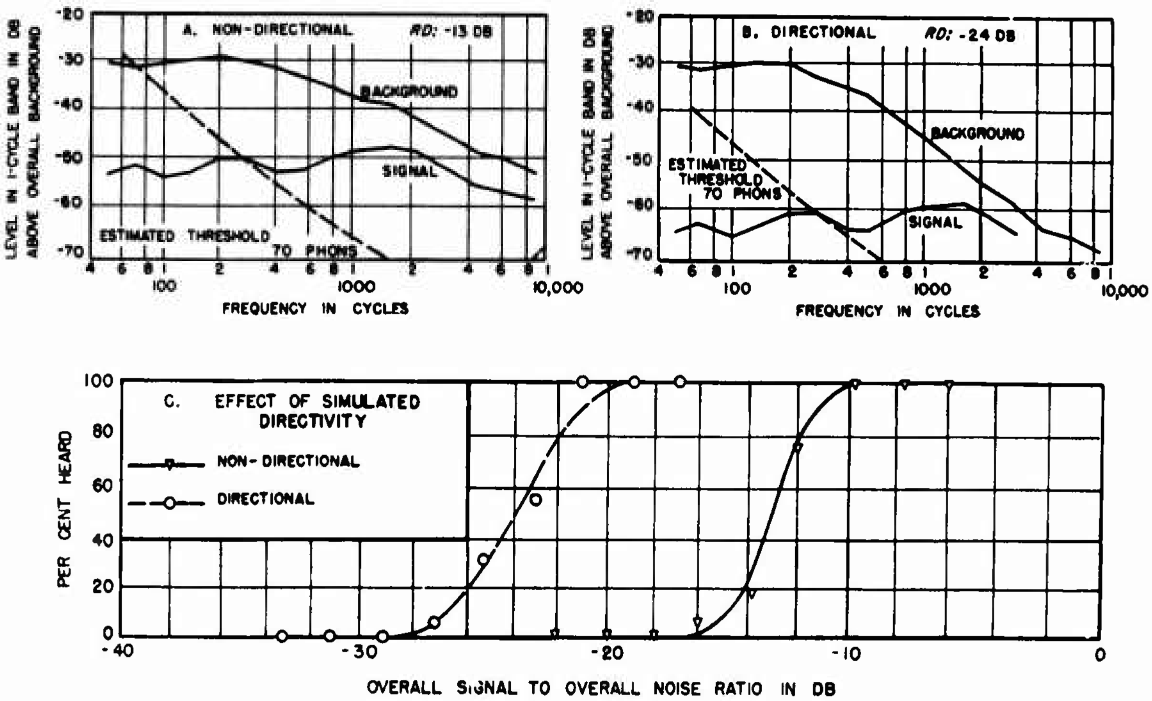


FIGURE 16. Audibility of sonic noise from both propellers at 200 rpm (4.5 to 5 knots), masked by water noise. Character of signal: propeller thrash.

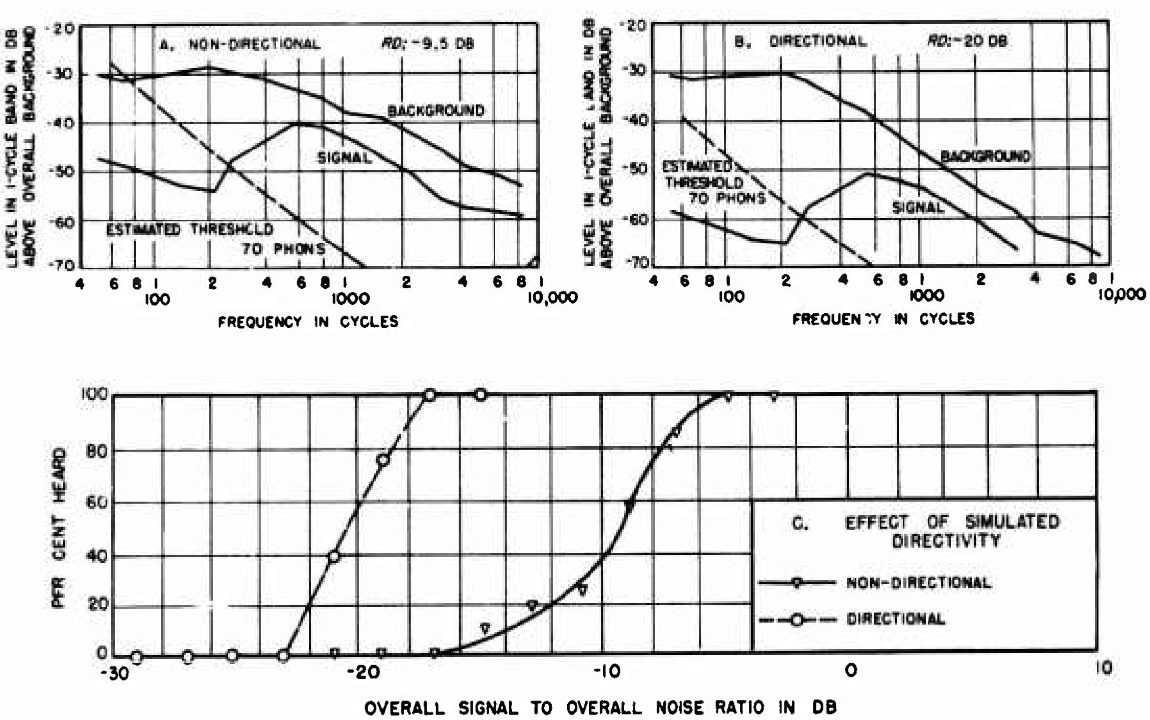


FIGURE 17. Audibility of sonic noise from one propeller at 200 rpm and periscope depth, masked by water noise. Character of signal: propeller thrash, but slower and more musical than the signal in Figure 16.

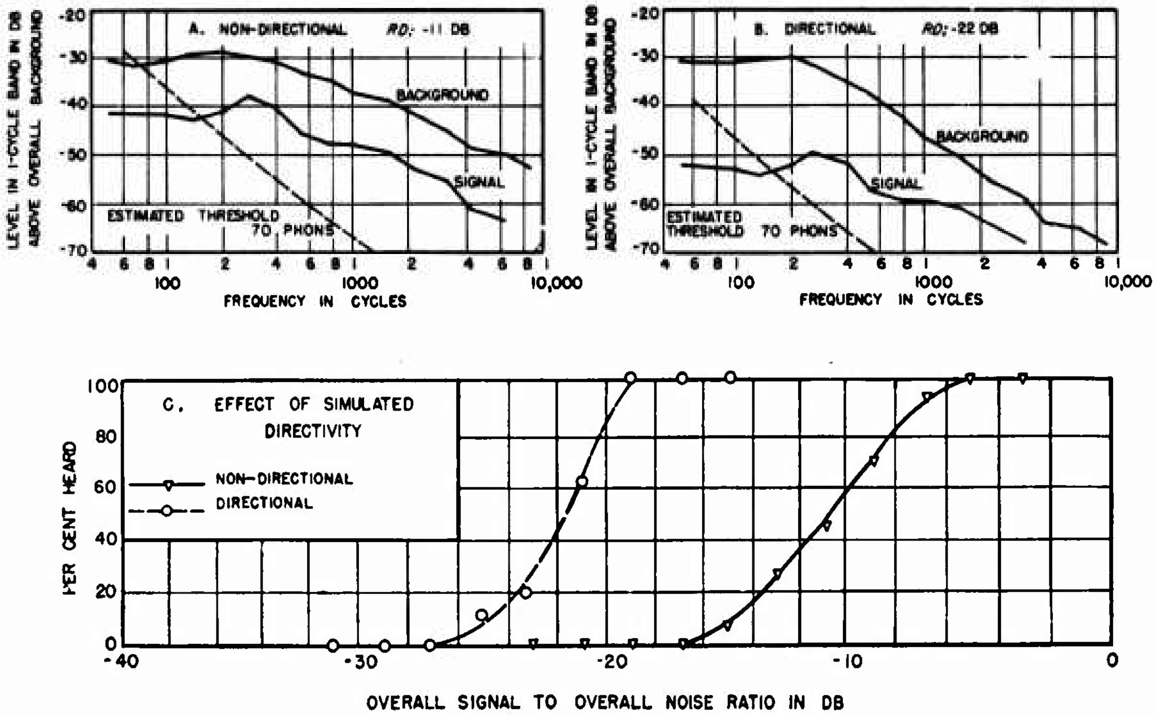


FIGURE 18. Audibility of sonic noise from main motors and propellers at 160 rpm (approximately 3.5 knots) and periscope depth, masked by water noise. Character of signal: propeller thrash and a steady hum of medium pitch.

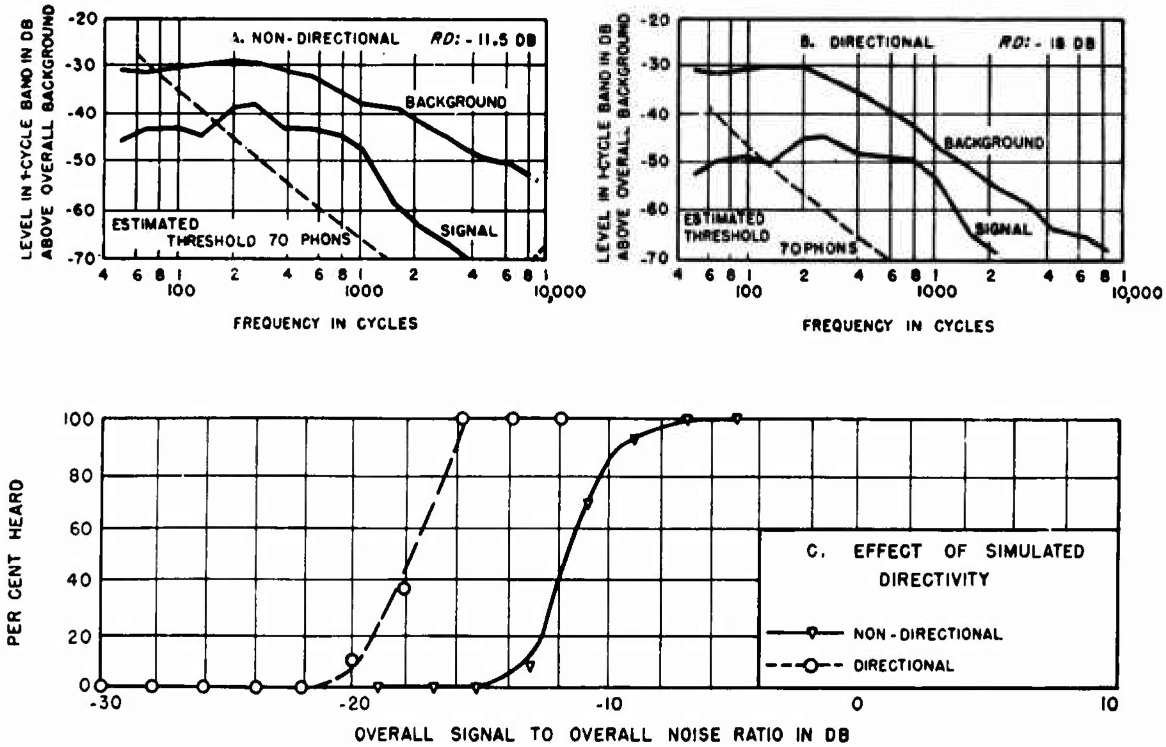


FIGURE 19. Audibility of sonic noise from gears at 90 rpm, masked by water noise. Character of signal: pulsating complex hum, together with medium and high pitched tones.

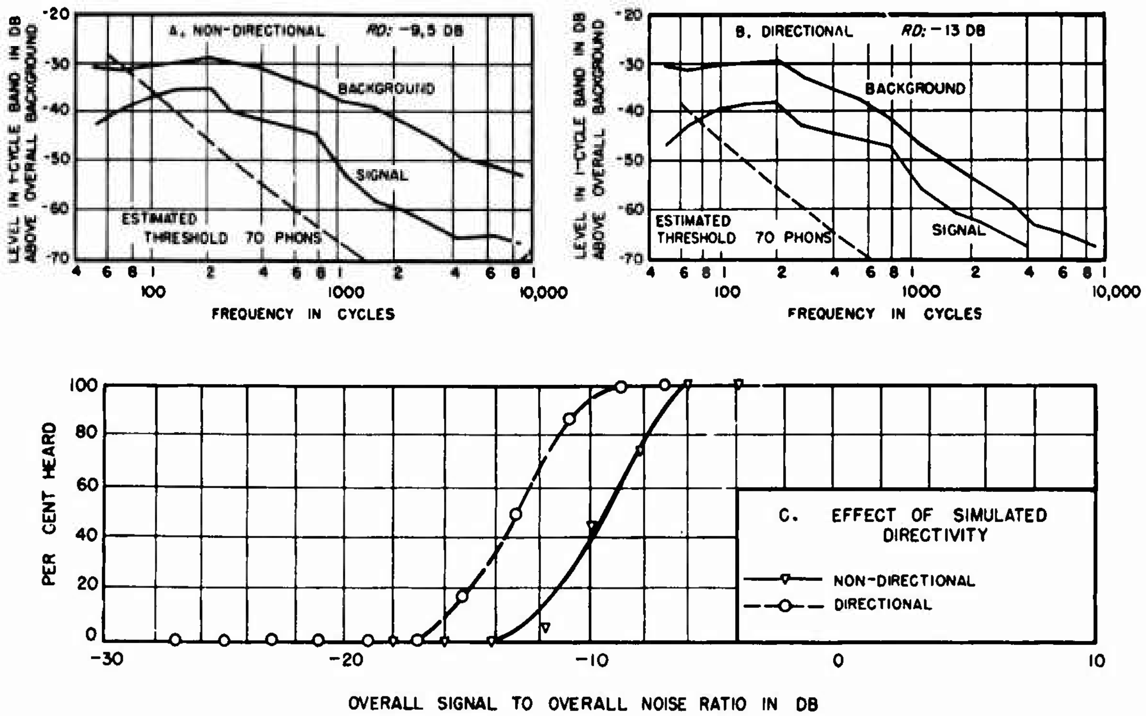


FIGURE 20. Audibility of sonic noise from gears at 60 to 70 rpm, masked by water noise. Character of signal: complex hum, fluctuating in pitch and intensity.

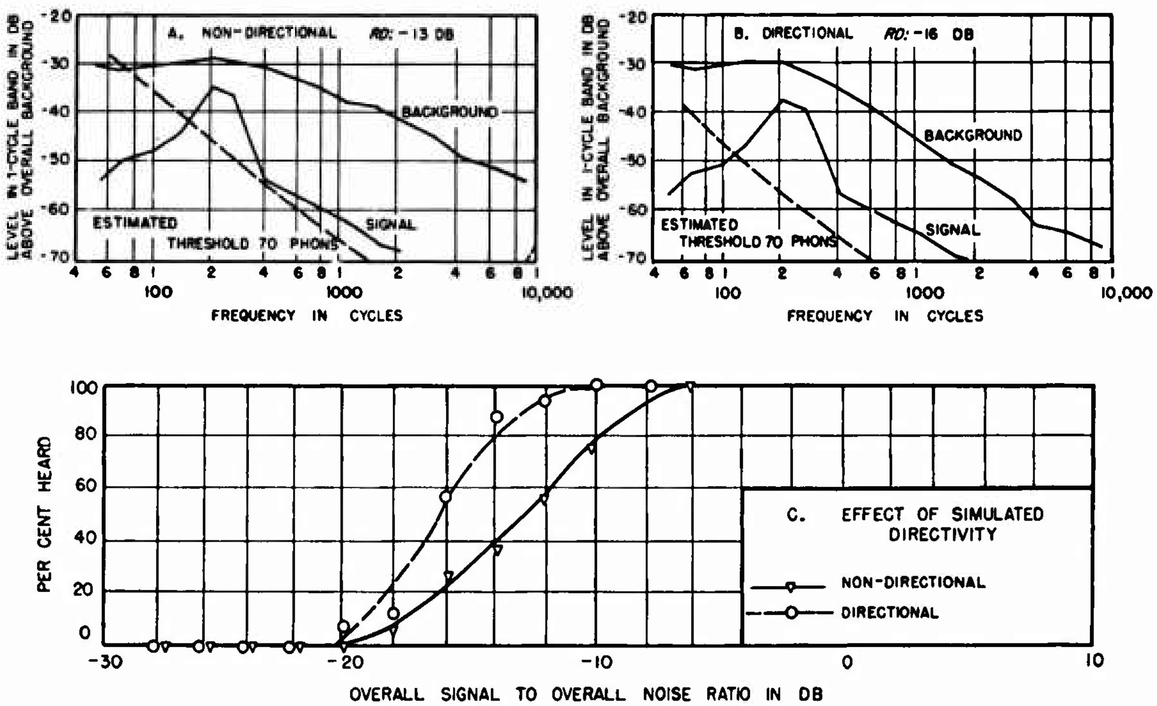


FIGURE 21. Audibility of sonic noise from the ballast pump, masked by water noise. Character of signal: steady sound of medium pitch.

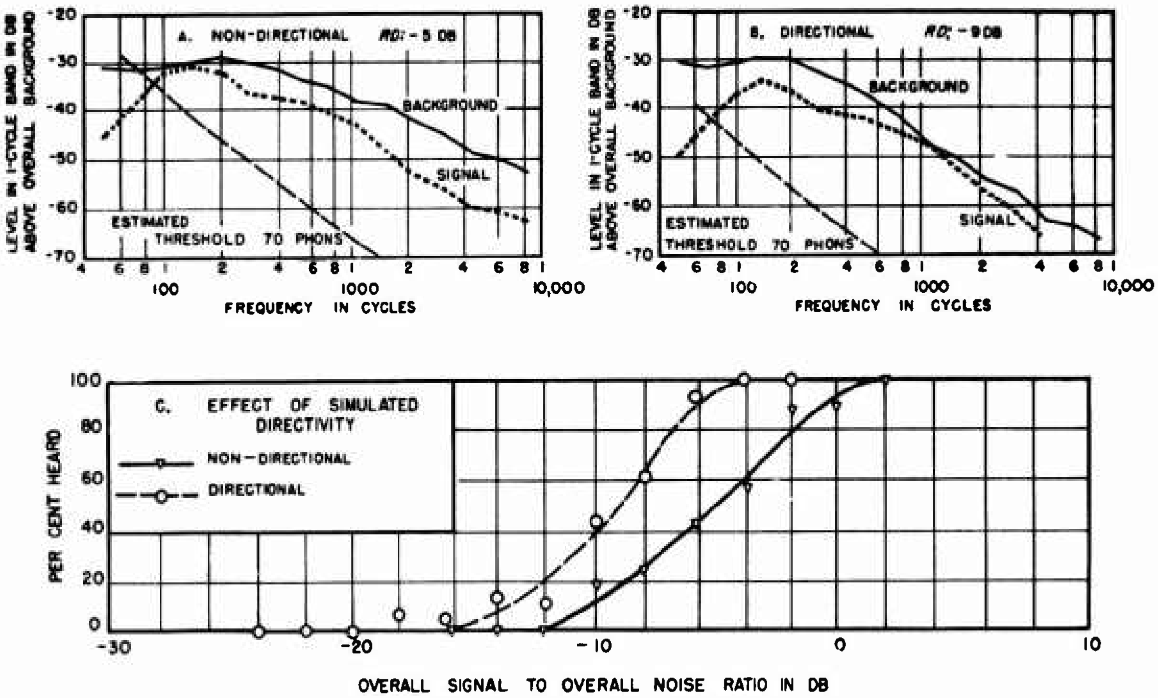


FIGURE 22. Audibility of sonic noise from circulating pumps at full speed, masked by water noise. Character of signal: faint hum, rushing, and gurgling noises, with no predominant characteristic.

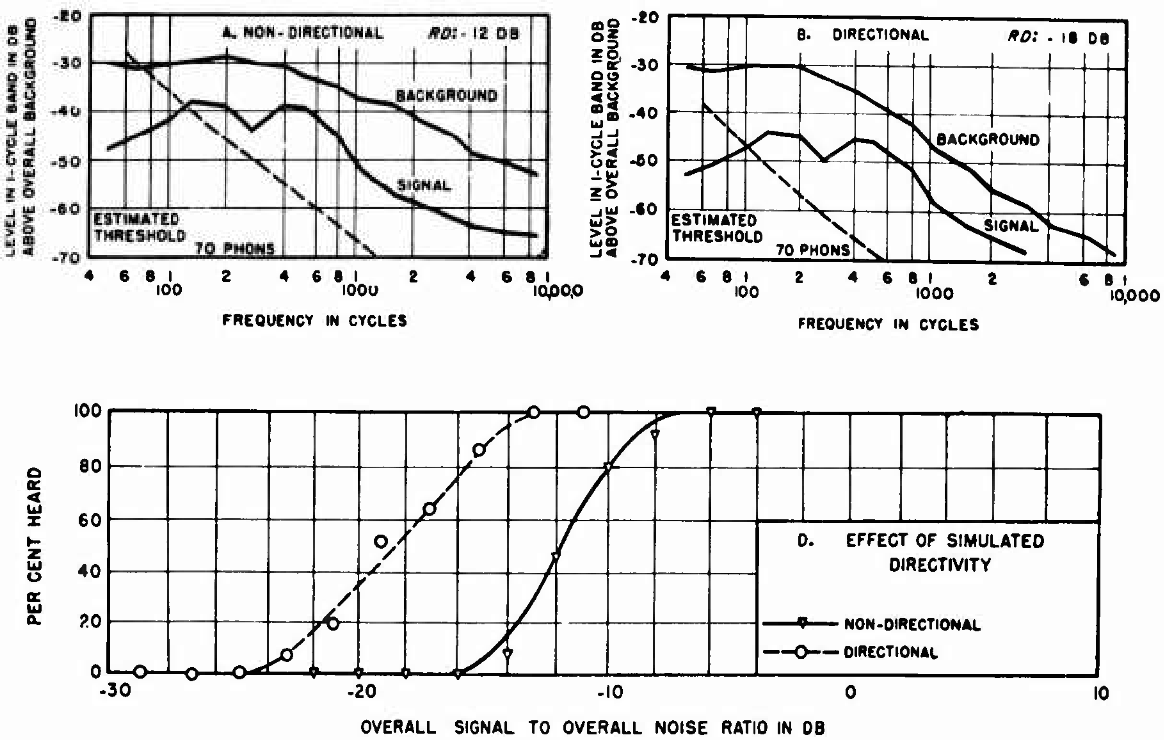


FIGURE 23. Audibility of sonic noise from circulating pumps at half speed, masked by water noise. Character of signal: steady hum of high pitch and gurgling noises.

proach. Use of the latter criterion instead of the *RD* reduces the variation from a factor of 1,000 (30 decibels) to a factor of 10. The following tabulation shows that the signal-to-noise ratios at the frequencies of closest approach are better criteria of signal audibility than are the recognition differentials referred to a wide band.

	Signal-to-noise ratio in db, in 0.1- to 10-kc band	Signal-to-noise ratio in db, at frequency of closest approach
Mean of 27 measurements	—9.5	—5.5
Average deviation from mean	4.3	2.5

These averages apply to all the spectra shown in Figure 6 through 23, except for the 5 high-pass filter cases. The latter are somewhat more difficult to interpret than the others; their inclusion would have little effect on the tabulated values. In other words, the ability to hear a distributed signal in the presence of a distributed background does not depend primarily on the total energy contents of the sounds. It is sufficient for detection that the energy content of the signal in an audible and relatively narrow frequency band approach the energy content of the background in that same frequency band. This amounts essentially to a restatement of the critical-band criterion for the audibility of tones in the presence of distributed backgrounds. Actually, it is an extension of that rule to the case in which the signal is not a pure tone, and in which both signal and background show a high degree of fluctuation, or intermittence. This extended critical-band rule implies, in addition, that the events in any one critical band are, to a first approximation, independent of the events in any other critical band, and that recognition of a broad-band signal may actually occur at a single optimal frequency or within a very narrow optimal frequency band. Of course, if the spectra of all underwater sounds had the same shape and time pattern, the recognition differentials would be identical for all signals in the presence of all backgrounds.

This concept is of fundamental importance. If it is an adequate statement of what determines the detectability of underwater sounds received in listening gear, it supplies the answer to a large number of operational and design problems. In addition, it furnishes a guide to the proper means of obtaining answers to a variety of still unsettled problems. To test the validity of this concept, a number of inferences from it may be examined; in other words, the status of its validity may be analyzed in terms of a small number of related questions which have been or can be settled experimentally. The general nature of these subsidiary questions and the answers thereto are outlined in the following paragraph and also at various points in the remainder of this section; a more extended discussion of the details will be found in Section 4.2.2.

Perhaps the first question which comes to mind is this: are optimal frequency bands, that is, groups of signal frequencies which are solely responsible for signal detection, actually observed? This seems to be the case, as shown in Section 4.2.1. Secondly, why, if the critical-band concept is valid, are the spectra separated by as much as 10 decibels at the optimal frequency (point of closest approach)? From the critical-band rule for tones, signal and background energies would be expected to be equal in the optimal critical band. There are at least two factors which contribute to this apparent departure from the critical-band rule: (1) fluctuation in signal and background levels and (2) obliteration of detail in the spectra caused by using filters which are wider than the critical bands. Thus, the peak signal levels were observed to rise 3 to 5 decibels above the mean levels which were used in determining the relation existing between signal and background spectra when the signal was just audible. It is probable that the signal was detected at these peak levels. In addition, if the levels of the various frequency regions in the spectra had been determined with narrower filters, these levels would probably be some 5 decibels higher than those shown in the neighborhood of the audible tones indicated in the various figures.

This estimate is reached in the following way. Suppose a pure tone of frequency f and intensity I_T is mixed with a wide-band distributed sound whose average intensity per cycle is I , and the energy of the mixture is measured through each of two filters with the same mid-frequency f, one of the filters having a pass band Δf_{cb} equal to the width of a critical band at the frequency f and the other having a pass band Δf_{oct} equal to the width of an octave centered at the frequency f. What would be the ratio of the "spectrum levels" of the mixture, as deduced from measurements with the two filters? The values of these spectrum levels may be obtained as follows.

1-cycle spectrum level obtained with a critical-band filter.

1a. Total intensity

$$I_{cb} = I_T + I_{\sim}\Delta f_{cb}$$

2a. Average intensity per cycle

$$\frac{I_{cb}}{\Delta f_{cb}} = \frac{I_T + I_{\sim}\Delta f_{cb}}{\Delta f_{cb}}$$

1-cycle spectrum level obtained with an octave filter.

1b. Total intensity

$$I_{oct} = I_T + I_{\sim}\Delta f_{oct}$$

2b. Average intensity per cycle

$$\frac{I_{oct}}{\Delta f_{oct}} = \frac{I_T + I_{\sim}\Delta f_{oct}}{\Delta f_{oct}}$$

By dividing 2a by 2b to obtain the ratio R of average intensities deduced from the narrow and wide filters and by denoting $(I_T/I_{\sim})$ by K, we obtain after simple manipulation,

$$10 \log R = 10 \log\left(\frac{K}{\Delta f_{cb}} + 1\right) - 10 \log\left(\frac{K}{\Delta f_{oct}} + 1\right). \quad (1)$$

This is the difference, in decibels, between the spectrum levels obtained by the two methods of measurement.

The value of this difference in level will be estimated for several representative cases. Consider a tone at 1 kilocycle, where $\Delta f_{cb} = 50$ cycles, and $\Delta f_{oct} = (2/3)\,1{,}000$ cycles. If the intensity of this tone were 100 times as great as that of the distributed sound in a 1-cycle band in the signal ($K = 100$), the tone would be an audible part of the signal when no background was present but would not be relatively strong. On the other hand, if the intensity of the tone were 1,000 times greater than that of the signal, in a 1-cycle band, the tone would be rated strong, when the signal was heard in the absence of the background. The value of 10 log R for these two cases is computed below (see also Figures 59 and 60).

$$10 \log R = 10 \log\left(\frac{100}{50} + 1\right) - 10 \log\left(\frac{100}{666} + 1\right)$$
$$= 4.3 \text{ db} \quad (2)$$

$$10 \log R = 10 \log\left(\frac{1000}{50} + 1\right) - 10 \log\left(\frac{1000}{666} + 1\right)$$
$$= 9.2 \text{ db.} \quad (3)$$

For an exceedingly strong tone, 10 log R would approach its limiting value 10 log $\Delta f_{oct}/\Delta f_{cb}$, amounting in this case to 11.2 db.

Hence, the presence of an audible tone at 1 kilocycle in the signal implies that the signal spectrum would in general be at least 4 decibels higher if the measurements had been performed with critical-band filters. The major effect produced by the use of the octave, rather than critical-band filters, is this reduction of level at the frequencies where tonal components occur; a correlated effect is the elevation of regions remote from the tone frequency, because several of the octave filters were able to pass the tone.

When the optimal component in the signal is a tone with a frequency in the neighborhood of 125 cycles, the width of the octave filter centered at the tone frequency is approximately equal to the width of the critical band stimulated by the tone; in other words, $\Delta f_{oct} = (2/3)$ 125 cycles, whereas Δf_{cb} is about 45 cycles in this frequency region. Under these circumstances octave filters give adequate resolution of the test sounds, and the tone-to-noise ratio should be approximately unity when the signal is just detectable. This situation is illustrated in Figures 7A, 9A, and 22A. When the tone has a somewhat higher frequency so that $\Delta f_{oct} > \Delta f_{cb}$, the spectra of signal and background tend to be spaced by some 5 to 10 decibels at the frequency of closest approach. These situations are illustrated in Figures 8A, 11A, 15A, and 20A.

In some cases, the masked signal becomes audible as a band of frequencies. The spectrum deduced from octave-band measurements then tends to be fairly comparable to what would be obtained from critical-band measurements and, provided the wide-band component responsible for signal detection does not exhibit a strong modulation, detection would be expected to occur when signal and background spectra coincide at the optimal frequency. This situation is illustrated in Figures 13A, 13B, and 22B, and also in Figures 7B, 8B, 11B, and 14B.

Finally, when the masked signal is detected as a wide band of frequencies, that is, when signal and background are parallel over a large frequency interval in the optimal region, and the optimal component has a strong and characteristic modulation, as in the case of propeller cavitation, detection depends essentially on the ability to discriminate changes of intensity rather than changes of quality. Under these conditions, recognition would be expected when signal and background spectra are spaced by 5 to 9 decibels at the frequency of closest approach. The basis of this estimate is discussed in Section 4.2.3; illustrative cases are shown in Figures 12A, 16A and 16B, 17A and 17B, and 18B. From these figures it will be noted that the observed spacing between signal and background is 6 to 10 decibels, for modulation rates in the neighborhood of 3 cycles, in excellent agreement with expectation.

The general conditions which must be met in order to assure aural detection of various types of sounds masked by specific kinds of noises are summarized in Chapter 6. It is worth noting some related practical matters at this point, since these aspects of the general problem do not depend essentially on the further experimental details given below. In the first place, it is obvious that the measurement of ship spectra at maximum detection distances is not feasible, since the signal levels, except in the optimal detection band, are usually far below background. Secondly, the submarine noise analyses given here are not entirely satisfactory from either a scientific or a practical standpoint, since they represent measurements on a small number of vessels, none of them current American fleet-type submarines, and the analyses were made with wider filters and slower recording instruments than seems desirable. Nevertheless, they are the only detailed submarine machinery spectra available at the present time. A few narrow-band measurements on the machinery sounds produced by fleet-type submarines have been made recently, but the results of these tests have not yet appeared in usable form. Another point which should be mentioned is this: enemy listening gear may be able to detect occasional, or highly intermittent sounds (see Figures 12 and 13) when they occur, but these may be inadequate for maintaining contact with a target.

4.1.4 Transition Curves

The process of signal detection and its relative efficiency are affected by a number of conditions which enter into the test situation and into the practical situation as well. In this series of tests, the following factors were varied, one at a time: the sequence in which the different signal levels were presented, the total loudness of the mixture, and the effect of filters. The influence of each of these factors is best shown by the group of transition curves which are included in the various figures. The various transition curves are shown together with the spectrum of the signal which was used in making the test, and the titles attached to the transition curves indicate the factor whose effect is depicted by the curve. The discussion of each of these variable factors, based on an examination of these transition curves, is given in the following four sections. Certain general considerations about transition curves are presented here first.

The transition curves, as their name implies, show that the emergence of recognizable differences among successive mixtures of sounds, as a function of their relative intensities, is not infinitely sharp. The probability that a signal will be detected increases gradually from nearly zero to nearly unity over a range of some 10 decibels. The standard procedure of taking the 50 per cent point as *the* signal-to-background ratio at which detection occurs is a reasonable approximation in many cases, but it is not equally appropriate to the analysis of all tactical situations. The practice has therefore been to define the precise detection probability which is intended in a particular discussion, for example, 50 per cent *RD*, 80 per cent *RD*.

For present purposes, it is sufficient to draw attention to the fact that the transition curves obtained in various tests are an important supplement to the information obtained from the spectra of the sounds used, and that they should be consulted along with the other relevant data if a detailed understanding of the detection process is desired. It is plain, therefore, that such curves should be described together with other results of listening studies.

The information supplied by a significant transition curve is the probability of perception under given circumstances. It is consequently important that the observers have the proper attitude, in other words, that they avoid guessing as well as excessive caution in reporting faint signals. Each of these kinds of bias tends to produce a characteristic distortion of the typical transition curve, since each deviates from the curve of perception by giving a preponderance of errors, either of omission or of commission. Thus, the inversion in the transition curve indicated by the circles in Figure 7D and the extremely gradual decrease of perception with decreasing signal level shown by the curve indicated by the circles in Figure 10C are distortions commonly due to guessing. Alternatively, an unusually gradual rate of decrease in detection probability with decreasing signal level may mean that some feature of the signal is particularly easy to recognize, and therefore detection is impaired less by a drop in signal-to-noise ratio than is ordinarily observed. Illustrations of a steeper than average slope, which often characterizes the curves given by excessively cautious observers, may be noted in Figures 13D and 19C.

Since guessing falsifies the results of listening tests, it is essential to eliminate all scores bearing evidence of this tendency, and, when observers show a persistent bias in this direction, to eliminate them from the tests. A standard device used to indicate this kind of unreliability is to include in the test a group of presentations which contain only the background sound; if an observer reports the signal when it is absent, his performance is unreliable. This procedure of including "blank" presentations was employed in the tests under discussion. The percentage of the blanks reported audible is shown by the symbols on the left-hand borders of those transition curves appended to Figures 6 through 15 which are labeled "effect of loudness" and "effect of filters." It will be observed that the number of false reports was vanishingly small except in the two cases already mentioned (Figures 7D and 10C). Even here, this tendency appears for only one condition of test, and it seems to have originated from confusion rather than recklessness (see Section 4.1.5).

There is no equally simple way of diagnosing an excessively cautious attitude on the part of the observer. In general, however, the transition curves given by such observers are significantly steeper than those obtained from "average" or "unbiased" observers taking the same test. Since transition curves obtained in laboratory tests represent performance under unusually good conditions, with distractions, boredom, and fatigue reduced to a minimum, it seems likely that a conservative attitude should be encouraged in the observers. This is particularly true since there is no laboratory equivalent to the serious or, at least, embarrassing consequences which may follow a false report in the practical situation. It is not intended to imply that sound operators are overly cautious (they are, in fact, trained and instructed to report all doubtful contacts, and to describe them as such), but rather that the observer in most laboratory tests has the advantage of knowing that audible signals *will* be contained in the series of sounds presented to him. He is on the alert for signals and has no hesitation in identifying them as such because he knows that any difference in the sounds presented must be due to the signal. He need not stop to consider whether the faint and often fleeting change in character which he hears originates aboard his own vessel, for example. Furthermore, to simplify the analysis of the data, most laboratory tests are neatly subdivided into a group of listening intervals each featuring either a signal or a blank. It is clear, therefore, that the type of laboratory test under discussion tends to define an upper limit of performance, and it is important that this limit should be established. But it is equally important that a parallel study be made of the performance of average sound operators under typical field conditions (see also Section 4.2.6).

There is one important aspect of these listening tests which cannot be adequately represented by the spectra, the time patterns, or the transition curves, namely, the subjective processes of perception, memory, and judgment which are associated with signal recognition. Some insight into these subjective matters can be gained through discussion with experienced observers, or better yet, by participating in a few typical listening tests.[d] For example, the degree of subjective certainty with which a signal can be identified as "audible" may often be fairly low. This is particularly likely to be true for faint signals (those with levels corresponding to less than 50 per cent detection probability) and for presentation methods in which the successive test sounds are widely spaced in time and not subject to control by the observer. Under these circumstances the observer must compare the sound just presented with a mental image of the masking background in order to decide whether any change has occurred due to the introduction of a signal; he may often pause for several seconds before arriving at a decision. As pointed out before, this decision may be influenced by the observer's personal bias. Listening under these conditions requires a high degree of attention and concentration; hence, a protracted interval of such testing is fatiguing and irritating and often yields highly variable results.

The problem of fatigue is met in practice by relieving the sound operator at fairly frequent intervals.[5] The particular procedure followed in this connection is probably somewhat variable; since no systematic study of optimal procedure has been made, present practice is presumably based on a mixture of expediency and practical experience. It has been suggested that a signal generator, controlled from the bridge and arranged to radiate an underwater signal of known strength to the receiving hydrophone, would be a useful aid in maintaining and checking the alertness of sound operators. Such a device should also be useful in helping to adjust the listening gear and to determine the extent of the masking background at different times and under various conditions of operation.

[d] Much of the material in the present section is described here for the first time. The discussion is based largely on still unpublished observations made in listening studies by University of California Division of War Research [UCDWR] and at the Mountain Lakes Station of CUDWR-USRL. The contents of these informal communications are introduced at the present point in order to unify the discussion of the problem.

Transition curves can in general be represented by an expression involving two parameters: (1) the signal-to-background ratio corresponding to 50 per cent detection probability, and (2) a quantity n which is related to the steepness of the transition curve. This basic expression may be written in the form

$$\frac{1-P}{P} = \left(\frac{S_{50}}{S}\right)^n, \qquad (4)$$

where P is detection probability, S is signal intensity, and S_{50} is the signal intensity corresponding to 50 per cent detection probability.

It may be noted that the shape and steepness of the available transition curves are not markedly different for fluctuating and nonfluctuating signals, when either is masked by a fluctuating wide-band background. Hence, equation (4) applies about equally well to both types of masking. For both types of curve, the mean value of n is approximately 3. Extreme values of n are about 1.5 and 9.

The exponent n is directly related to the "spread" of the transition curve, where this spread is defined as the number of decibels increase in signal-to-background ratio which is required to raise the detection probability from 20 per cent to 80 per cent. Substituting in equation (4) the pairs of values for P and S appropriate to 20 per cent and 80 per cent detection probability, respectively, gives

$$\frac{0.8}{0.2} = 4 = \left(\frac{S_{50}}{S_{20}}\right)^n, \qquad (5)$$

and

$$\frac{0.2}{0.8} = \frac{1}{4} = \left(\frac{S_{50}}{S_{80}}\right)^n. \qquad (6)$$

Dividing equation (5) by equation (6), and converting to decibel form, yields

$$10 \log 16 = n\left(10 \log \frac{S_{80}}{S_{20}}\right). \qquad (7)$$

The parenthesis in the right-hand member of equation (7) is the quantity defined as the spread in the previous text. Hence, finally, the spread is

$$\frac{10 \log 16}{n} = \left(\frac{12}{n}\right) \text{db}. \qquad (8)$$

Thus, the spread is 4 decibels when n equals 3. Furthermore, the spread diminishes, that is, the steepness increases, as n increases. A range of n between 1.5 and 9 corresponds to a range of spreads between 8 and 1.3 decibels. It is convenient to describe transition curves in terms of the spread as defined here because the numerical value of the spread is easily obtained by inspection; hence, n is readily evaluated with the aid of equation (8). Since also the "tails" of the transition curve (the intervals corresponding to detection probabilities of less than 20 per cent or more than 80 per cent) are usually not well defined, a purely observational quantity, such as the spread, is more appropriate than a theoretical parameter which defines the tails of the transition curve as well as the intermediate region.

When examining these transition curves, it should be borne in mind that they are composite curves obtained by lumping together the responses of all the observers in the listening group; slightly different results are obtained by averaging the number of signals detected by the entire group at a given signal-to-background ratio and by averaging the signal-to-background ratios required to give a particular detection probability for all observers. Furthermore, different observers taking the same test sometimes give transition curves with rather widely different values of n. In general, however, such differences tend to be minor. The various *RD* values obtained with a group of experienced observers rarely differ by more than 1 to 2 decibels.

The transition curves given by experienced observers are found to differ somewhat from those obtained with inexperienced observers, as the following data show. Ten observers were employed for that part of the tests under discussion, which are described in Sections 4.1.5 and 4.1.7. The five most consistent and reliable members of this group of ten were used in the remaining parts of the test program. The major difference between the transition curves obtained with the large and the small group was the greater tendency toward errors of commission shown by the less experienced observers; the cleaner "yes-no" transition characteristic of the probability curves in Figures 16 through 23 is typical of the more reliable observers. There was no significant difference in the *RD* values obtained with reliable observers and with observers inclined toward guessing. The mean difference in recognition differentials among the group of five consistent listeners was 0.2 decibel, although in one case (see Section 4.1.5) a member of this group failed to detect the signal when it was 6 decibels higher in level than the remaining four observers required for 50 per cent detection probability.

Several operational conclusions are indicated by the preceding discussion. The 50 per cent points are useful guides in computing general performance, but in cases where it is desired to estimate the signal levels corresponding to nearly certain detection, 6 decibels should be added to the 50 per cent signal-to-background ratio, and similarly, 6 decibels should be subtracted from the 50 per cent value to estimate the level of undetectable signals. If n is set equal to 3 in the transition curve equation and the value of the term 10 log S/S_{50} is either increased or decreased by 6 decibels, the corresponding values of detection probability are found to be 98.5 per cent and 1.5 per cent respectively. The spread in the transition curve indicates that, as the tactical situation varies from one condition to another, a detection probability greater or less than 50 per cent may be more significant than the average or 50 per cent detection probability. In addition, many targets may be first detected by means of prominent tonal components in their acoustic outputs. Such tonal components may not be sufficiently characteristic to permit identification of the target and will almost certainly be of no help in making a propeller turn count. In such cases, the range would have to be closed to obtain further information by means of listening.

4.1.5 Effect of Presentation Method

The effects of two different methods of presenting the test sounds were examined. In both these methods, the level of the masking background was maintained constant (after it had been adjusted to a loudness level of 70 phons), and the signal was introduced at various levels which differed by at least 2 decibels and covered the range from inaudible to audible. In both methods, also, the mixture of signal and background was presented for about 15 seconds, and a silent interval of 3 or 4 seconds separated successive presentations in order to allow time for recording signal audibility. The chief difference between the two methods of presentation was the fact that successive signal levels were presented in random order in one case, but in order of gradually increasing intensity in the other case. The random order method was intended to minimize the effects of guessing (and therefore included blank presentations); the gradually increasing level method was intended to simulate the field situation which might occur when the range to a submarine is gradually closed (and therefore included no blank presentations). The results are given in those transition curves in Figures 6 through 15 which are labeled "effect of presentation." These results seem to demonstrate, as indicated below, that there is no significant difference between the two methods (in the laboratory, and for wide-band listening). They demonstrate also that the observer should make a conscious effort to examine all the frequencies in the presentation band.[e]

Random Order Tests

Ten observers were used in the random order tests, which preceded the tests with increasing signal level; the background was maintained at a level of 70 phons. As preparation for the random order tests, the observers were presented with the background alone, with the signal alone, and finally were permitted to listen to the mixture while the signal gradually faded out. Each set of tests was performed twice, the order being changed in the duplicate test.

The transition curves give the average result. Since no mention is made of significant changes in performance on successive tests with a given signal, it may be assumed that practice effects were negligible. "It was found desirable," according to the report[1] describing the results of these tests, "not to make the order completely random, but to avoid the first level presented being a borderline case and to have an easily heard level early in the series. If the series begins with a number of 'not heard' observations the observers tend to be discouraged, and are more inclined to give random 'heard' observations at very low levels." This procedure seems a very poor approximation to the field situation, where the operator may fail to obtain a single contact through the entire course of a watch or, indeed, many consecutive watches. It is therefore desirable to simulate the practical condition in a small number of laboratory tests, and this has been done (see Section 4.2.6). Obviously, the expenditure of time required to conduct all tests in this manner would be prohibitive. Furthermore, it would be surprising if the change in testing procedure produced a large effect on the results. Nonetheless, it is useful to know what allowance, if any, should be made in applying laboratory results obtained by standard methods to the practical situation.

Increasing Signal-Level Method

The five most reliable observers were tested by the increasing signal level method. The results for these same observers by both methods of test are shown in the transition curves (Figures 6 through 15) labeled "effect of presentation." Comparison of the transition curves given by these five observers under the random method of test with the transition curves given by the entire group of ten observers for the random method, as shown in the curves labeled "effect of loudness," indicates that the five experienced observers gave recognition differentials which were not more than 0.4 decibel lower, on the average, than those obtained with the entire group of ten.

[e] It is convenient to distinguish between the *standard reference band* which extends from 0.1 to 10 kilocycles and the *presentation band* which is the group of frequencies actually presented to the listener. These frequency bands do not necessarily have the same limits.

In this method of test, the background alone was presented first. Then one of the signals previously used but not otherwise identified for the observers was introduced at a level which, from the earlier tests, was expected to be inaudible; the signal level was subsequently increased in steps of 2 decibels. The observers were requested to distinguish between certain and doubtful perceptions of the signal in order to check guessing. A maximum of two "possibly heard" reports was accepted, on the assumption that if an increase of 4 decibels in signal level was insufficient to produce a large increase in subjective certainty, the "possibly heard" report should have no weight. In general, little trouble was encountered from this source (the observers' knowledge that each presentation contained the signal) and the character of the transition curves obtained by both methods of test is substantially the same.

It is significant that in no case was the *RD* larger (or recognition more difficult) for the increasing-level method than for the random-order method and that in four of the eight cases studied by both methods there was a substantial improvement in the observers' ability to detect faint signals when the increasing level method was used. The evidence indicates that this improvement was real, that is, not owing to practice effects or other spurious causes. It also indicates that the improvement depended on alertness rather than on the presentation method *per se*.

Thus, in the random-order tests, the fact that the observers had foreknowledge of the character of the signal sound when it was not masked—either because it was presented alone or because it was presented together with the background but at a relatively high level — predisposed them to listen for the most characteristic features of the signal during the masking tests. Such foreknowledge is not typical of search conditions in the field. Furthermore, it is irrelevant and misleading in masking studies because the more striking aspects of a wide-band signal are not necessarily preserved when that signal is just audible in the presence of a masking background, inasmuch as the optimal component of the primaudible signal may not be a distinctive feature of the pure signal.

It was found during the course of these studies that if the signal were presented alone at a comfortable loudness, and then gradually faded out, with no masking background present, its character remained very much the same except for those signals containing a low-pitched hum. For these sounds the low-pitched component disappeared first, as would be anticipated from the fact that the sensation levels (number of decibels above the audibility threshold) at the lower frequencies are generally smaller than at the higher (see, for example, Figures 7, 9, 10, and 14). Hence, when the overall level of such a sound is diminished, the low-frequency components will be reduced to inaudibility before the higher frequencies fall below threshold. It was found also that, in the presence of the masking background, the character of the signal often changed drastically when its intensity was reduced toward the masked threshold while the background level was held constant. Those features which were most characteristic of the signal when heard alone were the first to be lost in the presence of the background, for example, the rapid knocking of the aft planes (Figure 6), the sharp crackle of the ballast pump (Figure 7), and the propeller thrash of the slow, submerged submarine (Figure 14). These are high-frequency sounds; thus, a masking background may impair signal audibility in a quite different frequency region from that affected by the auditory threshold. In all cases (except speech, Figure 13, and propeller thrash, Figures 16, 17, and 18) the primaudible signals were finally detected as more or less steady hums or tones; this was found to be the case with either presentation method.

Thus, the improvement obtained with the increasing-level method of presentation was apparently due to the fact that the observers were not listening most efficiently in this particular set of random-order tests. For efficient listening, it is necessary to search actively every part of the presentation band for traces of the signal, and not merely to attend passively to the entire presentation band or to a restricted frequency interval within which the signal is expected to occur. Active search was encouraged in the increasing-level tests by the fact that the character of the signal was not known to the observers in advance of the test. Table 1 compares the recognition differentials obtained with both presentation methods for the cases in which this factor made a significant difference. It will be noted that only four of the eight signals gave results which depended on presentation method, and that, when the observers learned to listen effectively, the improvement was maintained upon repetition of increasing-level trials. In other words, the change was not due to instability in the observers' reactions. However, if the random-order tests had followed the increasing-level tests, rather than preceding them, it is possible that a different result would have been obtained.

TABLE 1. Effect of presentation method.

Signal	Transition curve	Recognition differential in db				Approximate improvement in audibility in db
		Random-order tests	Increasing-level tests			
			1st trial	2nd trial	3rd trial	
1. Ballast pump	Figure 7C	−12	−12	−14	−15	2
2. Forward planes	Figure 8C	−19	−21	−21	....	2
3. Cooling fan	Figure 10B	−8	−9.5	−15	−13.5	6
4. Port propeller	Figure 14C	−9	−12	−15	−14	5

A given wide-band signal, masked by a given background, may be detected at different frequencies by different observers, or by the same observer at different times. That band of frequencies which permits detection of the faintest primaudible signal, in the presence of a particular background, is called the *optimal frequency band*. Failure to respond to the optimal signal frequency may require an increase of 6 to 8 decibels in the level of the primaudible signal (see Table 1). It is significant that the three signals (Figures 7, 10, and 14) for which there was the greatest variation between the results obtained with different presentation methods were the signals which also gave the greatest variation in results for individual observers. With the ballast pump signal (Figure 7), for example, one observer failed to detect the signal until it reached a level 6 decibels higher than that at which it had been detected by the other four. Thus, the use of a 6-decibel "safety factor," which was suggested at the end of Section 4.1.4 for the sake of assuring near certainty of detecting a given frequency component in the signal, seems warranted for the additional reason that a sound operator may overlook the optimal component. It seems clear from these observations that the cue which permits detection of the faintest signal does not necessarily thrust itself upon the attention. A diligent search may often be needed to find it, and failure to look for such a cue may be expected to yield results no better than would be obtained if the optimal cue were absent or suppressed.[f] Therefore the development of proper search habits should probably be emphasized in the training of sound operators. This need for careful examination of the received sounds, as well as random changes in the levels of signal and background, impose a lower limit on the time that an operator should allot to the output from a given hydrophone or from a hydrophone trained on a given bearing.

Failure of an observer to respond to the optimal cue, even after the signal-to-background ratio is substantially increased above the values

[f] It is interesting to compare the transition curves listed in Table 1 with those illustrating the effect of filters which attenuate the optimal frequency band (see Section 4.1.6). Careless search and actual suppression have essentially the same effect; the transition curves are shifted, with no significant change in shape, to higher values of the signal-to-background ratio.

required by other observers, may be due either to confusing similarity between the optimal cue and the background or to unfavorable presentation. Both factors seem to have played some part in the tests under discussion. For example, the ballast pump (Figure 7) and the cooling fans (Figure 10) are predominantly low-frequency signals, and were considered by the observers to be the most difficult to detect with certainty. Examination of the spectra for these cases indicates that the frequencies below 200 cycles were not far above the estimated threshold, whereas the background frequencies above this value were presented at considerably higher sensation levels. Hence remote masking of the optimal signal band, or at least distraction due to the background components at higher frequencies may easily have occurred. This phenomenon is especially likely because acoustic leakage around the headphone cap, with consequent reduction of intensity, is a significant problem for these lower frequencies. Furthermore, it was observed that the *RD* for the cooling fan was considerably smaller when the background was presented at a level of 90 phons; in other words, increasing the loudness of the low-frequency components improved the observers' ability to detect the optimal cue.

The spacing between signal and background spectra in Figures 10 and 14 would be increased by 5 to 6 decibels if the results of the increasing-level tests had been used rather than those of the random-order tests. However, the spectra were not analyzed with sufficient precision to permit any very significant conclusions to be drawn from this change.

4.1.6 Effect of Filters

Electrical filters were inserted between the mixer and the mixture amplifier (Figure 1), thereby passing or suppressing the same frequencies in signal and background. This procedure simulates the effect of filters in practical installations. Two types of filters were used: (1) a high-pass filter which transmitted all the frequencies higher than 1 kilocycle and which had a fairly sharp cutoff below that frequency and (2) various octave band-pass filters which were intended "to pass the predominant signal frequencies." The predominant frequencies were determined by examination of the signal analyses, plotted as energy per octave. Such a basis for choice is unsatisfactory for two reasons: (1) the ear's critical bands do not correspond to octaves, and (2) the optimal frequency band depends on the relative levels of signal and background spectra, and not primarily on the absolute level at various frequencies in the signal spectrum. Despite these objections the filter tests provide useful information.

These tests were carried out with the signals and background shown in Figures 6 through 15. The random-order method of presentation was followed, and the same group of five most reliable observers was used as in tests previously described. All the tests were made at a listening level of 70 phons. Since the filters excluded large parts of the standard reference band from the ears of the observers, the sensation levels at the transmitted frequencies were higher than in the tests in which the entire 0.1- to 10-kilocycle band was presented at a level of 70 phons. To the extent to which auditory acuity improves with increasing sensation level, the presentation of the optimal frequencies at a higher sensation level should permit the detection of fainter signals, and the results seem to show that this factor played a small but noticeable role in this group of tests.

Table 2 lists the improvement in detectability produced by the insertion of the various filters. Improvement is here defined as the difference between the *RD* without filter and the *RD* with filter. Thus, if the *RD* with filter is a larger negative number than the *RD* without filter, the difference is positive; in other words, there is an improvement or increased ability to detect weaker signals. Conversely, a negative difference indicates a deterioration due to the use of the filter. The recognition differentials (decibel difference between primaudible signal and background in the standard reference band at the input to the filter) are used rather than the presentation differentials (decibel difference between primaudible signal and background at the output of the filter) because performance in the field is best expressed in terms of the sounds-in-the-water rather than the sounds-at-the-ear. Similarly, the pertinent transition curves, which are listed in Table 2, all show the signal-to-background ratio at the inputs to the filters. All these transition curves represent the performance of the same five observers for various conditions of test.

loss brought about by failure to make a careful search for the optimal cue in this particular signal (line 4 of Table 1). It is possible, therefore, that the optimal component in the case of the port propeller occurred in the neighborhood of 200 cycles. The primaudible signal, in the case of the ballast pump, was difficult to detect with certainty because its character was similar to that of the background. If this similarity were accentuated by the use of the filter, some deterioration would be expected. The probability that such an effect occurred is indicated by the greater number of errors of commission observed when using the band-pass filters (see transition curves labeled "effect of filters"). The slight improvement shown for most of the other cases is of the order of magnitude of experimental error; if the improvement is real, however, the effective narrowing of the critical bands with increased sensation level would be sufficient to account for it.

TABLE 2. Effect of filters.

Signal	Transition curve	Band-pass filters		High-pass filter	
		Passband in cycles	Improvement in audibility in db	Transmitted spectra at primaudibility	Improvement in audibility in db
1. Aft planes	Figure 6E	220–440	0	Figure 6B	−9
2. Ballast pump	Figure 7E	110–220	−2	Figure 7B	−14
3. Forward planes	Figure 8E	165–330	2	Figure 8B	−24
4. Auxiliary circulator	Figure 9D	440–880	1	....	−2
5. Cooling fan	Figure 10D	110–220	1	....	Not detected at any level
6. Lubricating oil separator	Figure 11E	110–220	1	Figure 11B	−14
7. Aft periscope	Figure 12C	2,400–4,800	1	....	0
8. Speech	Figure 13D	1,300–2,600	0	....	0
9. Port propeller	Figure 14E	440–880	−4	Figure 14B	−3
10. Main motors	Figure 15D	660–1,320	0	....	−2

The third column of Table 2 gives the limits of the octave filters which were used in these tests because they passed the "predominant" signal frequencies. These same limits are indicated by the horizontal arrows in the 0.1- to 10-kilocycle spectra shown in Figures 6 through 15; the portions of the spectra contained within these limits give a fair picture of the sounds presented to the ear, since the change in signal-to-background ratio due to use of the filters is quite small for all the signals except the port propeller (see the fourth column of Table 2). Examination of the spectra indicates that the optimal octave band was not always selected (see Figures 9A and 14A). This circumstance apparently had little effect in the first case (line 4 of Table 2) but produced a deterioration of 4 decibels in the second (line 9). The available data provides no information concerning the source of the difference in results, but it is interesting to note that the deterioration of 4 decibels shown in line 9 of Table 2 is very nearly equal to the

Column 6 of Table 2 shows the effect of removing the low frequencies by inserting a 1-kilocycle high-pass filter. For those signals in which the optimal octave band was below 1 kilocycle (column three of Table 2), the high-pass filter hampered detection; in those cases in which the detection frequency was above 1 kilocycle, the high-pass filter tests gave exactly the same recognition differentials as were obtained without the filter, showing that the presence of the lower frequencies did not mask the higher. In other words, listening systems which do not admit the low frequencies are at a disadvantage in detecting machinery sounds. This deterioration amounted to 24 decibels in the case of the forward planes, whereas in the case of the cooling fan, the recorded signal contained nothing but background noise above 1 kilocycle.

Column 5 of Table 2 lists several figures showing the relation between signal and background at primaudibility when the 1-kilocycle high-pass filter was employed. The filtered spectra are shown for only those cases in which use of the high-pass filter produced a large change in *RD*. It will be observed from these figures that a considerable amount of energy was passed between 500 and 1,000 cycles; it is possible, therefore, that signal recognition actually occurred below 1 kilocycle, despite the large slope of the spectra in that region. To check the latter possibility, it would be necessary to explore the sounds admitted by the high-pass filter with the aid of a narrow band-pass filter.

The near identity of the recognition differentials obtained without filters and with band-pass filters admitting the optimal signal component, and the subjective similarity of the primaudible component heard under both conditions of test, imply that the optimal frequency is generally the recognition frequency. Exceptions may be due to (1) inattentiveness, (2) elimination of the optimal frequency through insertion of filters, as described previously, or through inadequacy of system band width, and (3) poor response of the detector in the optimal region. When the ear is the detector, frequency translation may be needed in order to give the optimal frequency favorable presentation in those cases where it lies beyond the limits of the standard reference band.

4.1.7 Effect of Loudness Level

These tests were conducted to determine whether the recognition differential is significantly affected by the absolute level of the background; in other words, by receiver gain. The level of the background was set at 50, 70, or 90 phons and maintained at one of these settings while the signals shown in Figures 6 through 15 were presented to ten observers by the random-order method. The transition curves obtained are labeled "effect of loudness," and the fractional number of errors of commission in the blank presentations is indicated on the left-hand borders of the figures. The spectra drawn in Figures 6 through 15 may not apply equally well to all three loudness conditions, since frequency and amplitude distortion are difficult to eliminate. It is probable, however, that such distortion is not smaller in most practical installations than in the system used for these tests.

The 70-phon level was rated most satisfactory by the observers. The 50-phon setting necessitated holding the breath while listening and would consequently invite interference from room noise at most listening locations in the field, whereas the 90-phon setting was considered too fatiguing for prolonged listening periods. In addition to these general effects, some dependence of *RD* on listening level would be expected because of the operation of such factors as threshold limitation of the optimal component, improvement in pitch and intensity discrimination with increased sensation level, and changes in the degree of remote as well as adjacent masking (see Figure 5 in this chapter and Figure 2 in Chapter 2). The effects of these factors depend on the frequency of the optimal component and the compositions and time patterns of the presented sounds. The net result may be either an improvement or a deterioration in performance.

In all but one case (Figure 10C), the observed effect of loudness level on ease of signal detection was fairly small (2 to 3 decibels or less). It was found that at the 90-phon setting the low-frequency components tended to dominate the presented sounds, as would be anticipated from the steeper rate of increase in loudness with sensation level for this region (see Figure 5 and Section 2.2.2). This effect probably enhanced the listener's ability to discern the low-frequency optimal component indicated in Figure 10. Thus, the improvement of 7 decibels shown by Figure 10C to result from increasing the gain is almost exactly equal to the improvement produced by careful listening (see Figure 10B and line 3 of Table 1).

Figure 5 implies that threshold-limited audibility of optimal components was not a likely contingency for the signals shown in Figures 6 through 15, even at the 50-phon level, except for frequencies below about 140 cycles per second. The only signal for which such an effect might have been expected is shown in Figure 7. However, the observed effect of loudness level in this case was not significantly greater than for most of the other sounds studied. It may be noted in this connection that the various trials summarized in Figure 9 indicate that the low-pitched hum at 100 cycles played no part in the detection of this signal under masking-limited conditions. Since the octave and critical bands centered at 100 cycles have very nearly the same width, the 100-cycle component was actually fainter relative to background than was the 400-cycle hum specified as the primaudible component in Figure 9. Threshold limiting of the optimal component shown in Figure 11 is improbable because it rose in pitch during the course of the record.

In two of the ten cases examined (Figures 8D and 9C) loudness had no effect, either beneficial or adverse, upon signal detection. In five cases (Figures 6D, 7D, 10C, 11D, and 14D) performance improved progressively with increased loudness of the mixture, although errors of commission showed a parallel increase. Although the number of false reports assumed significant proportions for only one of these signals (Figure 10C), the trend implies the operation of an unfavorable factor. This factor may be increased distortion in the reproducing system, and, if so, can probably be eliminated by improved design. It is interesting to note that the optimal component in these five cases seems to have occurred below 300 cycles.

In one of the remaining three cases (Figure 12B) the 90-phon level resulted in somewhat poorer scores than the 70-phon level; in the other two (Figures 13C and 15D) performance deteriorated progressively with increasing overall gain. For these three cases, the optimal component seems to have occurred at or above 800 cycles. Although remote masking must be considered as a possible explanation of the adverse effect of increased gain for these three signals, the effect should be offset by use of high-pass filters, but this did not occur for the cases in question (lines 7, 8, and 10 of Table 2). It seems more likely, therefore, that distortions introduced by the test apparatus influenced the results.

Further study is required before detailed conclusions can be drawn, but the available information appears to warrant several general remarks. To begin with, the gain setting should be such that the masking background is heard at comfortable loudness. Since the frequency of the optimal component is generally unpredictable, this procedure favors audibility of the widest possible band of frequencies without undue risk of operator fatigue or distortions produced by overload of the listening system. Because of the unpredictable frequency of the optimal component during search, no simple selectivity scheme is available for emphasizing or improving the presentation of a narrow detection band; all the components in the presentation band contribute to the loudness and thereby limit the sensation level which may be obtained within the detection band. In the second place, auditory acuity tends to improve with increasing sensation level even in the presence of masking.

4.1.8 Effect of Simulated Hydrophone Directivity

The purpose of these tests was to investigate the improvement in audibility to be expected from use of directional hydrophones capable of discriminating against nondirectional ambient-noise backgrounds. Since such hydrophones also give some discrimination against the high-frequency components in the target sounds, two filters were used to simulate the directional condition: (1) a background filter (with the discrimination characteristic shown in Figure 24) which was inserted between the pre-amplifier of the ambient-noise sound head and the mixer and (2) a signal filter (with the discrimination characteristic shown in Figure 24) which was inserted at the corresponding point in the signal channel. Figure 25 represents the difference between the curves in Figure 24 and thus indicates the improvement in signal-to-background ratio at various frequencies which is afforded by the directional properties of the particular hydrophone, an 8-foot ring, used as a model for these tests.

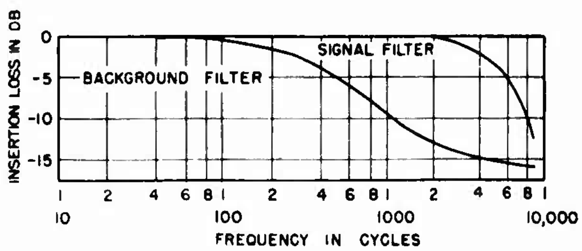


FIGURE 24. Effect of signal and background filters. The response of the latter was designed to simulate the discrimination against nondirectional ambient of an 8-foot ring hydrophone.

The discrimination shown in Figure 24 applies only to conditions in which the target acts as an extended source. Since the signals used in these tests represent quiet submarines, detectable only at fairly short range, the design of the signal filter is appropriate.

In the directional condition, the reproducing system simulated the sounds which would be received in a highly directional hydrophone trained on-target, and not subject to control by the operator; in the nondirectional condition, the system simulated the sounds which would be received from the same target, in the same prevailing background, through a nondirectional hydrophone. The sounds were presented by the increasing level method, and at a level of 70 phons, to five reliable observers, four of whom had participated in all the tests previously described. The spectra of these sounds, as presented in each condition of test, are shown in Figures 16 through 23. The resulting transition curves, which are given in the same figures, represent the average of two determinations for each condition; and the horizontal shift between the curves marked "with" and "without" filters shows the improvement in audibility, or change in *RD*. The recognition differentials, as well as the levels of overall background indicated on the ordinates, refer to sound levels at the inputs to the filters, that is, they correspond to sound-in-the-water. The pertinent observations are summarized in Table 3 and are discussed below.

Column 3 lists the estimated frequencies of the signal components detected in the directional and nondirectional conditions of test. This estimate has been made by noting the optimal frequencies, that is, the frequencies

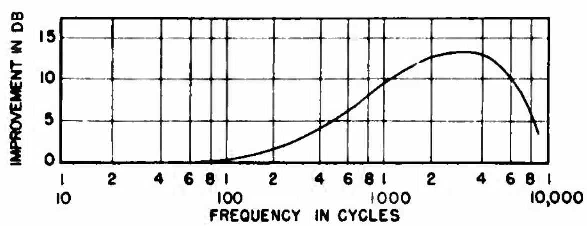


FIGURE 25. Improvement in signal-to-water noise ratio produced by filters.

corresponding to the distance of closest approach in the spectra shown in Figures 16 through 23. In column 4 are tabulated the spacings between signal and background spectra at the optimal frequencies. It will be observed that, while this spacing may be as small as 0 decibel, it averaged about 6 to 7 decibels, in agreement with the discussion in Section 4.2.3. In column 5 are shown the various amounts of improvement in detectability of the signals due to simulated hydrophone directivity. These improvements have been read from the transition curves, and, for comparison with them, the sixth column indicates the computed improvement.

When hydrophone directivity is introduced, the shapes of the spectra representing the sounds-at-the-ear are changed, as shown in Figures 16 through 23. Thus, two types of cases must be distinguished when computing the improvement in detectability resulting from directivity: (1) those in which the same signal component is recognized in both conditions of test and (2) those in which a different signal component becomes primaudible in the directional and nondirectional conditions. Examination of the spectra indicates that the detection frequency was not changed by the simulated directivity in six of the eight cases studied but was changed in the remaining two (see the fifth and seventh signals in Table 3 and also the corresponding spectra). The cases in which the frequency of the optimal component varied are discussed separately from the others.

essential parallelism of the signal and background spectra at the higher frequencies in Figures 16 through 18. Figure 25 shows that the maximum objective improvement, occurring at about 3 kilocycles, amounted to 13

TABLE 3. Effect of hydrophone directivity.

Signal	Transmitted spectra at primaudibility	Estimated detection frequency in cycles	Spacing between spectra at optimal frequency in db	Observed improvement of *RD* in db	Computed improvement of *RD* in db
1. Propellers at 200 rpm	Figure 16A	Broad band above 2 kc	6	11	13
	Figure 16B*	Broad band above 2 kc	6	..	..
2. Propeller at 200 rpm	Figure 17A	Broad band above 1 kc	6	11	13
	Figure 17B*	Broad band above 1 kc	7	..	..
3. Motors and propellers at 160 rpm	Figure 18A	Broad band above 1 kc	10	11	13
	Figure 18B*	Broad band above 1 kc	10	..	..
4. Gear noise at 90 rpm	Figure 19A	800	8	6.5	8
	Figure 19B*	800	7	..	..
5. Gear noise at 60 to 70 rpm	Figure 20A	200	7	3.5	3
	Figure 20B*	800	5	..	..
6. Ballast pump	Figure 21A	200	6	3.5	2
	Figure 21B*	200	7	..	..
7. Circulating pumps at full speed	Figure 22A	150	0	4	5
	Figure 22B*	1,100	0	..	..
8. Circulating pumps at half speed	Figure 23A	540	7	6.5	5.5
	Figure 23B*	540	8	..	..

*Filters were inserted to simulate the effects of hydrophone directivity.

As explained in the following paragraph, when the detection frequency is the same for both conditions of test, the improvement in *RD* should be numerically equal to the objective improvement in signal-to-background ratio afforded at the optimal frequency by the directivity filters. The various amounts of objective improvement for all signals other than 5 and 7 have been computed from column 3 and Figure 25 and are entered in column 6. In determining the computed improvement from Figure 25, the maximum possible improvement was selected in each instance, on the assumption that the signal would become primaudible as soon as the required detection ratio was attained in the optimal frequency band. Thus, lines 1, 2, and 3 represent situations in which the signals probably became primaudible within a fairly wide band of frequencies, as indicated by the decibels, but that a possible improvement equal to or greater than 10 decibels characterized the entire band between 1 and 6 kilocycles. The presumptive wide-band quality of the primaudible component in these cases probably accounts, in part, for the fact that the observed improvement fell somewhat short of the computed maximum improvement, which is associated with a narrower frequency band. Similarly, two distinct signal components, at 0.3 and above 1.0 kilocycle in Figure 18A, and at 0.3 and 0.8 kilocycle in Figure 19A could have and possibly did become primaudible in the nondirectional condition; here again, the maximum possible objective improvement (associated with the higher frequency) has been entered in column seven for the reason stated above.

The procedure outlined for computing the improvement in *RD* may appear curious at first glance, since the recognition differentials give the signal-to-background ratio in terms of the *overall* intensities of the sounds, whereas the objective improvements (as read from column 3 and Figure 25) apply to *specific* frequency regions. This procedure is not contradictory, however, since the directivity filter diminishes the level of the masking background at the optimal frequency relative to the overall level of the unfiltered background, which serves as the reference standard for both conditions of test. When the intensity of the optimal signal component is reduced a corresponding amount during a listening test, it remains primaudible, but all the other components in the signal are reduced comparably. Thus, the *RD* (or overall signal-to-background ratio at primaudibility) in the directional condition will be less than that in the nondirectional condition by an amount equal to the objective improvement in signal-to-noise ratio at the recognition frequency.

This improvement is given by Figure 25 and includes the effect of the directivity filter in the signal channel. The overall signal-to-overall noise ratio given by this calculation is that at the input to the filters (and corresponds to the sound-in-the-water, or *RD*, rather than that at the output, or presentation differential). In this particular case, the two types of differentials would differ by a nearly constant factor amounting to 3 to 4 decibels. This circumstance arises from the fact that the bulk of the energy, in sounds with large negative spectrum slope, is concentrated in the lower frequencies, whereas the directivity filters affect primarily the upper frequencies. The first part of the preceding statement may be demonstrated by replotting the filtered and also the unfiltered background spectra depicted in Figures 16 through 23 on scales linear with respect to energy and frequency. The ordinates will then represent energy per cycle and the abscissas, cycles. The integrated areas under each of the curves, obtained by counting squares or by planimetry, will represent the overall energies of the filtered and unfiltered sounds. The difference between them is of the order of 4 decibels for the masking background, as also demonstrated by sound level measurements during the tests. The overall levels of the signals, on the other hand, were hardly modified by the directivity filter in the signal channel, since that filter affected the highest frequencies only, and these contained a negligible fraction of the total energy.

The agreement between observed and computed changes in *RD* shown in columns 5 and 6 serves to verify two points emphasized in the preceding and ensuing discussions: (1) recognition in the case of broad-band sounds depends, in general, not on the total energies, but on the signal-to-noise ratio in a narrow frequency band; and (2) the optimal frequency is usually the recognition frequency. This reemphasis is, in fact, one of the chief reasons for computing column six. Another reason is the desirability of assessing the factors which determine the practical effects of hydrophone directivity. As indicated in Section 4.1.3, the numerical values of wide-band recognition differentials (and hence, their calculation) have little intrinsic interest. Indeed, the available information tends to support the view that ultimately operational prediction and analysis may be performed most satisfactorily by direct comparison of signal and background spectra, without explicit consideration of overall levels (see Section 4.2.2).

It is worth noting at this point that the criterion adopted for predicting the improvements in recognition differentials will be valid only if the change in the shapes of the spectra does not alter the signal-to-noise ratio required for primaudibility in the detection band. Such changes in the detection ratio might be produced by the modified sensation level at which the primaudible component is identified, or by differences in the degree of remote masking, of time pattern, or of distortion introduced by the test apparatus. In practice, hydrophone directivity may also change the peakiness of the received background and thereby the signal strength needed for primaudibility (see Figure 3 in Chapter 5). As indicated by column 4 in Table 3, the detection ratio for a given signal was essentially independent of the simulated directivity. This observation appears quite reasonable for the cases in which the same signal component was detected in each condition of test but is probably sheer accident for the cases in which different components were identified in the directional and nondirectional tests.

In the case of lines 5 and 7, the improvement also depends on the effectiveness of the directivity filters in reducing the level of background but the degree of improvement is modified by the fact that the frequency and character of the primaudible component are variable rather than fixed. Thus, for the fifth signal, the optimal component occurs at 200 cycles in the nondirectional condition and at 800 cycles in the directional. These results are in agreement with the approximate prediction of theory (see Section 4.2.2). Examination of Figure 20A shows that, because of the relative shapes and positions of the spectra, the spacing between nondirectional background and the (undetected) component at 800 cycles amounted to 10 decibels. In the second condition of test (Figure 20B) the spacing to the (detected) 800-cycle component is reduced to 5 decibels; in other words, this component, and therefore the entire signal, must be raised by 5 decibels to render it detectable, thereby losing 5 of the potential 8-decibel improvement shown by Figure 25 to accrue from directivity for a signal component primaudible at 800 cycles. The net gain, 3 decibels, has been entered in column 6. Thus, the improvement is greater than would have been observed had the 200-cycle component been detected in both conditions, but less than would be obtained for a fixed 800-cycle component. While the corresponding gain in range for such cases will in general have an intermediate value, as illustrated here, bearing accuracy will be improved in proportion to the change in frequency of the optimal component. The computed improvement for the seventh signal was found by this same method.

Several conclusions are worth noting at this point. From the last two columns in Table 3, it follows that knowledge of the signal and background spectra, together with the discrimination of the hydrophone, makes it possible to predict improvement in performance to within about 2 decibels. As a corollary, it follows that the hydrophone whose discrimination is shown in Figure 24 should give better general performance than the line hydrophone with the characteristic illustrated in Figure 18 of Chapter 3, provided the discrimination of hydrophones against nondirectional ambient noise is also an adequate index of their discrimination against self-noise. From the spacings shown in column four, it may be concluded that increasing the slope of the background spectrum from 6 decibels per octave (nondirectional condition) to 10 decibels per octave (directional condition) has no adverse effect on signal audibility. In fact, the loudness contributions from the various components in the directional background spectrum are more nearly equal, in the band between 0.2 and 6 kilocycles, than is the case for the nondirectional background (see Figures 16 through 23). Protracted listening to the directional background therefore is probably less fatiguing, and search for the optimal component may be somewhat easier, since the various frequencies receive more nearly equal emphasis at the ear and the high frequencies are less annoying. To assure effective listening over the widest usable frequency band, the outputs of practical systems should probably tend to parallel the shape of the audibility threshold. It will be observed that the relation between the 70-phon background spectrum and the estimated thresholds for distributed sounds is somewhat different for the directional and nondirectional cases because of the changed composition of the presented sound. Finally, the improvement produced by the directional condition is greater for the propeller cavitation sounds in Table 3 than for the machinery sounds since cavitation was detected at the higher frequencies, where the simulated discrimination was greatest.

4.2 UCDWR TESTS

A thorough and comprehensive program of listening tests, in progress at the time of writing at the San Diego Laboratory of the University of California Division of War Research [UCDWR] was undertaken to obtain fundamental information about the auditory detection of underwater sounds. The methods and results of this study, some of which have been made available in private communications prior to publication and some of which have appeared in published reports,[6,8] are described in Sections 4.2.1 through 4.2.6.

4.2.1 Techniques

The apparatus used in these tests is shown schematically in Figure 1. Care was taken to provide practically distortionless gain in the amplifier components. The recorded sounds which were used are identified in the various figures and tables presented below. The selections studied were obtained with a variety of sonic and supersonic receivers and include signals from ships, submarines, and torpedoes; in a few cases, simulated propeller sounds were employed. The sonic backgrounds included ambient noise, with and without shrimp crackle, and also simulated deep-sea ambient noise; the supersonic backgrounds were self-noise and ambient noise, with and without shrimp crackle. The simulated ambient was obtained from a broad-band, thermal noise source with a spectrum slope of −6 decibels per octave; simulated propeller sounds were obtained by similar means and were given various rates and degrees of amplitude modulation. Combinations of signal and background pairs from this library of recordings were presented through headphones worn by the observers. In general, the properties of the receivers and the techniques of recording and playback were such that a moderately faithful replica of the sounds-in-the-water was presented to the headphones. Since not all the available recordings were of high quality (their merit is indicated in the various tables, the rating being based on quality in a 0.1- to 10-kilocycle band and on freedom from extraneous noise) and since not all signal and noise pairs were obtained with the same hydrophone, the observations are not equally typical of the field situation. Thus, the peaks at 2 and 5 kilocycles in the spectrum of the recorded shrimp crackle (see Figure 28) are artifacts introduced by resonances in the hydrophone used for recording. More recently, higher quality recordings have been obtained by transmitting the signals over an FM link from the receiving vessel to a shore station, where the recording can be done under better conditions than are usually available on board ship. Also, apparatus for monitoring ship sounds has been installed at the entrance to San Diego harbor. The hydrophones used in this installation are mounted on the harbor bottom; their outputs are transmitted by cable to a shore station, where a frequency translator makes it possible to bring 8-kilocycle bands of supersonic frequencies into the audio region. Thus, high quality recordings of ship sounds, in the interval from 0.1 to 30 kilocycles, can be obtained.

The five observers who participated in each test were either members of the laboratory staff or enlisted men, newly graduated from the United States Navy West Coast Sound School. Since no individual in the latter group was available for more than a week, at least one experienced observer from the laboratory personnel was included in each test in order to serve as a control. The weekly addition of new observers tends to make the results obtained more representative and permitted study of such factors as learning and the effects of selection and Sound School training.

Various presentation methods were studied. Only the results obtained by means of random-order presentations are given in the tables and figures, the effect of other presentation methods being outlined briefly in Section 4.2.6. Random-order tests lasted between 20 and 30 minutes and were begun by permitting the listeners to familiarize themselves with the signals and backgrounds which were to be mixed in the subsequent test. The background was maintained at a constant level corresponding to between 60 and 75 phons, the gain applied to the mixture being subject to some control by the individual listeners, and the mixtures were presented at 10-second intervals for periods of 5 seconds. Signal level was varied in steps of 2 decibels and randomizing was achieved by means of an automatic selector. This selector controls the gain setting in the signal channel according to a predetermined pattern which is punched upon a moving tape. Several tapes were used in order to eliminate learning of the presentation order. If an observer reported that the signal was audible in more than one of the blank presentations, the entire test for that individual was rejected.

A recently developed device for recording listener response consists of a sound range recorder modified to give a group of parallel traces, which show at a glance the gain setting in the signal channel at any time during a test and the simultaneous response of the observers. The latter indicate their judgments by throwing a hand switch from a neutral position to either of two positions corresponding to "audible" and "inaudible." This procedure saves time and confusion in recording judgments and scoring performance; it indicates indecision and time needed to reach a decision; and it gives a compact and permanent record of simultaneous signal levels and responses. The latter feature permits elimination of the unrealistic subdivision of the test into a group of discrete listening intervals each of which contains either a signal or a blank. This subdivision is necessary in the usual test procedure in order to allow time for writing down responses and also to simplify the process of scoring the test, which involves setting up a correspondence between items in the response series and in the signal level series. Thus, it is relatively simple by means of the procedure just outlined to score performance during long intervals of continuous listening designed to simulate the shipboard situation.

"Following every test[7] each listener's description of the character of the primaudible signal was recorded. The same signal was presented to all five listeners but usually each described it in different words. Also the octave or half-octave band, necessary to pass the frequency component for making the signal primaudible, was determined by listening while a variable set of half-octave filters was inserted into the signal channel. It was generally found that the primaudible components of the signal were located in a limited frequency band less than an octave in width. In the usual case, where it was apparent from the listeners' comments that all had detected the same component of the signal, it was sufficient to determine the primaudible frequency only once." Care was taken, in obtaining these comments, to have the listeners report the character of the signal at primaudibility, and not the more striking feature of the sound when the signal was several decibels above primaudibility (see also Section 4.1.5). Occasionally, the primaudible component extended over a relatively wide band (see Figures 38 and 55).

The overall levels in a 0.1- to 10-kilocycle band of all the sounds used were measured, and the spectra were obtained by means of a 50-cycle band-pass filter. Therefore tonal components appear in the spectra as peaks 50 cycles wide (see also Figures 59 and 60). Power was measured with a level recorder of high writing speed so that time patterns could be studied. Time patterns for a number of typical cases are illustrated in many figures in this chapter. The power level traces represent independent measurements, within the recognition band, made in the noise and signal channels when the gain settings were adjusted to the ratio which rendered the signal primaudible. The time-intensity-frequency traces, described in Section 3.1.1, were made with a 45-cycle filter and with the aid of a tilting network which increased the gain of the higher frequencies at the rate of 6 decibels per octave, in order to confine the effective range of transmitted power levels to the contrast range of the recording paper. The rms spectra in the figures represent the ratio of power in a 50-cycle band to the overall power in the standard reference band and are plotted as a function of the midfrequencies of the 50-cycle bands. Such spectra will be termed *50-cycle spectra;* in the range 0.1 to 1.5 kilocycles they correspond very closely to *critical-band spectra* (power measured in frequency intervals of width equal to a critical band; see Figures 76 through 79). At the higher frequencies, however, the 50-cycle spectra differ considerably from critical-band spectra, and this must be borne in mind when interpreting the observations. In order to investigate the presence of tonal components, comparisons were made of the power passed by half-octave and 50-cycle filters in the frequency region where such components appeared to exist. The significance of the arrows and the shaded areas in these diagrams is explained in the next section.

4.2.2 Observed Recognition Differentials

The results obtained for sonic target sounds are given in Table 4, which is subdivided into three sections, according to the background used in the tests. The relation between signal and noise spectra at primaudibility is shown for some typical cases in Figures 26 through 53, and the legends in these figures refer to the corresponding data in Table 4. These primaudibility charts, as well as the others in this volume, give, not the sums of signal and background levels, but independent spectra for each of the sounds, referred to the same over-all level. Hence signal-to-noise ratios may be obtained by inspection. As pointed out before, it was found that a wide-band acoustic signal generally becomes primaudible within a restricted frequency interval. These recognition frequencies are indicated in columns five and ten of Table 4. Alternatively, they may be read directly from the spectra by noting the frequency at which signal and noise curves are closest to each other (see, for example, the component at 360 cycles in Figure 26).

This close approach of signal and noise spectra at primaudibility implies that the critical-band criterion may properly be extended to complex sounds of the kind illustrated, since that criterion requires that signal and noise energies approach unity in at least one critical band. That the various critical bands are, to a first approximation, independent of each other is indicated by the fact that simultaneous, equal stimulation of two critical bands does not necessarily improve recognition (see Figure 34). There are some indications, however, that simultaneous stimulation of multiple critical bands may affect recognition (for example, see Section 4.2.3).

It was observed, when the components in the optimal 50-cycle band displayed a strong rhythmic fluctuation, that the average value of the rms signal level within that 50-cycle band could be as much as 12 decibels below the rms noise level at primaudibility (see lines 26 and 33 of Table 4, and Figures 40, 41, and 42). In all such cases, however, fairly consistent agreement with the critical-band criterion was obtained when the signal-to-noise ratio in the optimal critical band was computed from the recurrent peak values of the rms signal level. This is illustrated (1) in columns 8 and 9 of Table 4; (2) in the primaudibility charts in Figures 26 through 53, where the arrows and shaded areas signify fluctuating tones and fluctuating frequency bands, respectively; and (3) in the power level traces (included among Figures 26 through 53) which show the temporal relation between signal and noise levels in the optimal 50-cycle band at primaudibility. It will be noted from Table 4, however, that even the recurrent peak level of the signal may fall 2 to 3 decibels below the mean noise level, both measured in the same critical band. This may imply that fluctuating sounds are somewhat more audible than steady ones, due to the effects of auditory motion (see Sections 5.1 and 5.6), or that the measured level of a rapidly fluctuating component cannot build up to full strength when restrictive filters are used (see Section 4.2.3). These points may be worth investigating for the cases of frequency- and amplitude-modulated tones and may have some bearing on noise reduction measurements as well as on the choice of proper operating conditions for a vessel engaged in evasive maneuvers.

It follows, therefore, that signal-to-noise ratios at primaudibility may be predicted directly from a knowledge of the spectra involved, provided the frequency analyses are made with adequate resolution (that is, analyzer band width not exceeding critical band width) and provided also the dynamic response of the recorder approximates that of the ear. Since the latter has a build-up time of between 0.1 and 0.2 second, it can react rapidly enough to respond to the maximum levels reached by a pulsating signal modulated at 10 cycles or less.

To test the effects of recorder speed, traces were made using a rapid writing speed of 400 decibels per second and a slow speed of 100 decibels per second. In measurements of random noise, such as deep-sea ambient or shrimp noise, the average level recorded was practically independent of writing speed. When measurements are made of rapidly modulated signals like those illustrated in Figures 39 and 45, the rhythmic peak levels are slightly closer

TABLE 4. Primaudibility of sonic sounds.

Signal	Figure	Merit of signal	Character of primaudible signal	Recognition frequency region in cycles	Signal-to-noise ratio in db in a 0.1–10-kc band		Signal-to-noise ratio in db in the optimal 50-cycle band at primaudibility		Midfrequency in cycles of optimal 50-cycle band
					80 per cent recognition probability	50 per cent recognition probability	Average ratio	Ratio for rhythm peaks	
A. Background: simulated deep-sea noise.									
1. Anchored S-class submarine, charging batteries with diesels	36	Good	Low-frequency rhythm	140–280	−4	−6	3	....	120
2. Anchored S-class submarine, charging batteries with diesels	....	Fair	Low-frequency rumble	....	8	−10	0	....	400
3. S-class submarine at 5 knots and periscope depth	....	Fair	Rhythmic swish	....	10	−12	−2	....	900
4. Submerged S-class submarine at 6 knots	....	Fair	Rhythmic churning	....	9	11	−2	....	520
5. S-class submarine at 3 knots and periscope depth	....	Good	Rhythmic chug	200–400	8	9	−2	....	250
6. Submerged submarine at 5 knots	....	Fair	Rhythmic churning	....	−11	−13	−5	0	450
7. Fleet-type submarine at 5 knots and periscope depth	29	Good	Squeaking	Above 3,200	−12	−14	....	8	7,000
		Good	Rhythmic hiss	Above 3,200	−12	−14	−7	0	7,000
8. Fleet-type submarine at 60 rpm (3 knots) and periscope depth	26	Good	Hum or whine	280–400	−9	−12	0	....	360
9. Surfaced fleet-type submarine underway at 12 knots	34	Good	Tinny grinding	1,130–1,600	−4	−6	−1	....	1,400
		Good	Low rumble	140–280	−4	−6	−2	....	120
10. German Type VII C submarine (GRAF) at 2½ knots and periscope depth	50	Fair	Low bubbling rhythm	200–400	−9	−11	−8	2	300
11. Convoy (13 merchant vessels, 6 large tank lighters, and 3 PC escorts)	53	Fair	Hum or whine	....	−10	−12	−3	0	390
12. Tanker, 1 mile away	44	Fair	Low rhythmic noise, whine	565–800	−7	−9	0	....	680
13. Freighter, 7,000 tons, 442 feet long, at 10 knots	38	Good	Rhythmic swish and slow rhythm	Above 4,500	−8	−12	−1	1	8,600
14. Destroyer, 2,870 tons, 381 feet long, at 15 knots	43	Good	Low tone with propeller rhythm	....	−9	−11	−2	....	250
15. Torpedo (3-turbine type), at 45 knots	17	Good	High-pitched whine or hum	800–1,130	−9	−11	6	....	970
16. Torpedo (electric type), at 30 knots	49	Good	Hum and whine	....	−15	−17	−1	....	600

TABLE 4. (*Continued*)

Signal	Figure	Merit of signal	Character of primaudible signal	Recognition frequency region in cycles	Signal-to-noise ratio in db in a 0.1–10 kc band: 80 per cent recognition probability	Signal-to-noise ratio in db in a 0.1–10 kc band: 50 per cent recognition probability	Signal-to-noise ratio in db in the optimal 50-cycle band at primaudibility: Average ratio	Signal-to-noise ratio in db in the optimal 50-cycle band at primaudibility: Ratio for rhythm peaks	Midfrequency in cycles of optimal 50-cycle band
			B. Background: shrimp crackle. Merit: very good.						
17. Anchored S-class submarine, charging batteries with diesels	37	Good	Low-frequency rhythm	140–200	–8	–10	4	....	150
18. Anchored S-class submarine, charging batteries with diesels	....	Fair	Rhythmic rumble	280–565	–18	–20	–2	....	400
19. Submerged S-class submarine at 6 knots	....	Fair	Rhythmic churning	400–800	–16	–18	–2	0	550
20. Submerged S-class submarine at 5 knots	....	Fair	Rhythmic churning, whine	400–800	–11	–13	2	....	540
21. S-class submarine at 3 knots and periscope depth	....	Good	Rumble	200–280	–11	–15	–1	....	250
22. Fleet-type submarine at 5 knots and periscope depth	32	Good	Whine	400–565	–13	–15	0	....	500
23. Fleet-type submarine at 60 rpm (3 knots) and periscope depth	28	Good	Hum or whirr	280–400	19	–21	0	....	360
24. Surfaced fleet-type submarine underway at 12 knots	35	Good	Low rumble	140–280	–4	–7	6	....	120
		Good	Whirr	1,130–1,600	–4	–7	–3	....	1,400
25. Convoy (13 merchant vessels, 6 large tank lighters, 3 PC escorts)	....	Fair	Rhythmic hum	280–565	–20	–21	–3	0	390
26. Freighter, 7,000 tons, 442 feet long, at 10 knots	40	Good	Rhythmic chug	Below 400	–12	–14	–6	–1	260
27. Torpedo (3-turbine type), at 45 knots	....	Good	Whine	800–1,130	–19	–21	0	....	950
			C. Background: water noise recorded at a depth of 300 feet; water depth 1,000 feet.						
28. S-class submarine, at 6 knots and periscope depth	....	Fair	Rhythmic chugging	....	–9	–11	–5	–2	500
29. Submerged S-class submarine at 5 knots	....	Fair	Pulsating hum	....	–11	–13	0	....	360
30. S-class submarine at 3 knots and periscope depth	....	Good	Low rumble	....	–8	–10	2	....	250
31. Convoy (13 merchant vessels, 6 large tank lighters, 3 PC escorts)	....	Fair	Rhythmic churning	....	–10	–12	–5	–2	390
32. Tanker, 1 mile away	....	Fair	Rhythmic chugging	....	–9	–11	4	....	190
33. Freighter, 7,000 tons, 442 feet long, at 10 knots	....	Good	Rhythmic churning	....	–18	–20	–12	–3	180
34. Torpedo (3-turbine type), at 45 knots	....	Good	Tone	....	–10	–12	0	....	200
35. Torpedo (3-turbine type), at 45 knots	....	Good	Whine	....	–10	–12	–4	....	950

to the noise levels in the optimal critical bands when a fast writing speed is used. With slowly modulated signals, however, virtually any writing speed is satisfactory.

The following is a summary of the results listed in Table 4.

	Signal-to-noise ratio in db in 0.1-10 kc band for a 50 per cent recognition probability	Average signal-to-noise ratio in db in 50-cycle band at primaudibility	Signal-to noise ratio-in db in 50-cycle band for rhythm peaks at primaudibility
Mean (35 measurements)	−12.9	−1.5	0.0
Average deviation from mean	3.1	2.7	1.6

It is evident that the ratio of maximum signal-to-average noise, both measured in the optimal critical band, gives the most consistent

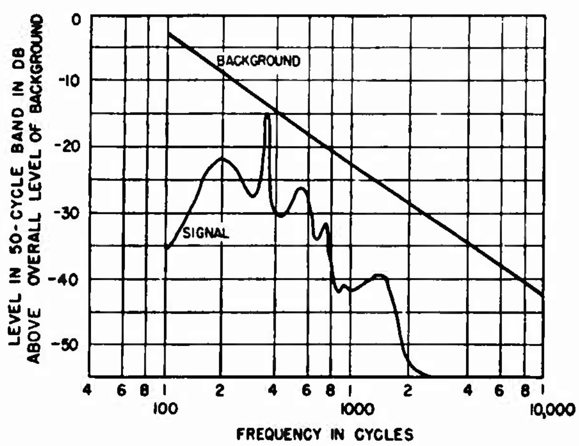


FIGURE 26. Audibility of sonic noise from a fleet-type submarine at 3 knots and periscope depth, masked by simulated deep-sea noise. *RD* for presentation band shown: —12 decibels. Character of primaudible signal: hum or whine. Primaudible frequency: 360 cycles. See Table 4, line 8.

prediction of the condition for primaudibility. The mean value of −1.5 decibels obtained when using the average signal level in the optimal 50-cycle band is probably not a reliable index of the amount of modulation which is typical of ship and submarine signals. Only a few of the primaudible components studied were amplitude modulated, and some of these fluctuated by as much as 15 to 20 decibels. Similarly, the comparatively small average deviation of 3.1 decibels among the wide-band recognition differentials is not necessarily typical, since the magnitude of the scatter depends on the degree of similarity of signal and noise spectra, on time patterns, and phase and harmonic relations. Consequently, no universal index, such as the recognition differential in terms of overall power, can have useful predictive reliability.

An analyzer band width of 50 cycles was employed through the range 0.1 to 10 kilocycles. This approximates the critical band widths in the region below about 1.5 kilocycles (see Figure 17 in Chapter 2) but is narrower than the critical bands for the higher frequencies. The 50-cycle band width was generally satisfactory, however, because few of the sounds studied contained prominent tones above 1 to 2 kilocycles. Strictly speaking, a critical-band analysis can be performed only with a group of discrete filters of fixed midfrequency and band width. This is complicated, experimentally. It was found satisfactory to use instead a heterodyne analyzer with a 50-cycle band width of variable midfrequency. Obviously, the error introduced by failure to use the correct critical band width will be the same for signal and background when these are distributed sounds with approximately the same slope in the region measured. Hence, this error cancels out when the difference between signal and noise levels is computed. Specific correction to the proper band width must be made, however, when tonal components are present above 1.5 kilocycles. In the case of distributed sounds, the 50-cycle spectra are uniformly 17 decibels higher in level than the 1-cycle spectra (10 log 50 = 17 decibels) and have the same slope as the latter (see also Figures 59 and 60).

Figure 29 represents a situation in which the signal was recognized as a high-pitched squeak, primaudible at 7 kilocycles. The optimal signal-to-noise ratio at primaudibility, in this case, was +8 decibels when signal and noise were measured with the 50-cycle filter (see line 7 in Table 4). Since there is a promi-

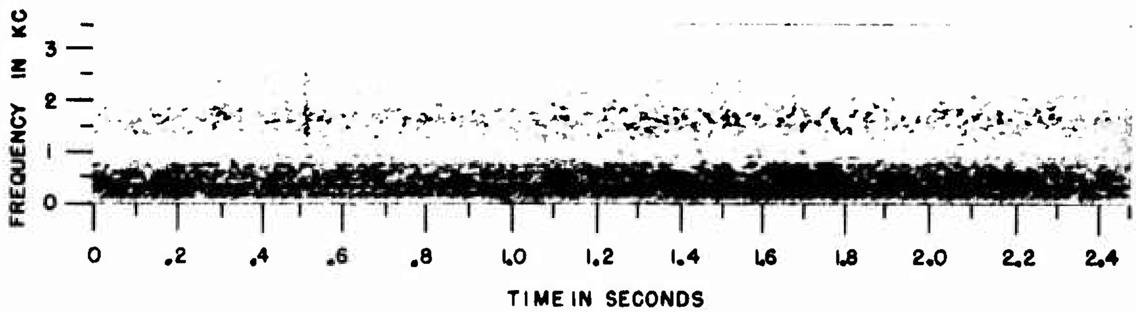


FIGURE 27. Time-frequency-intensity analysis for sonic noise from a 3-knot submarine (Figure 26).

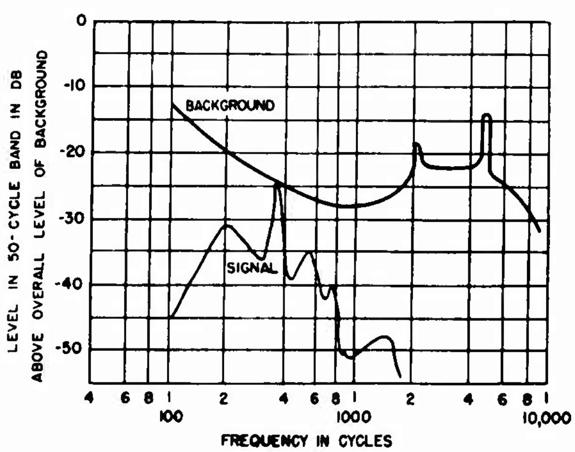


FIGURE 28. Audibility of sonic noise from a fleet-type submarine at 3 knots and periscope depth, masked by shrimp noise. *RD* for presentation band shown: —21 decibels. Character of primaudible signal: hum or whirr. Primaudible frequency: 360 cycles. See Table 4, line 23.

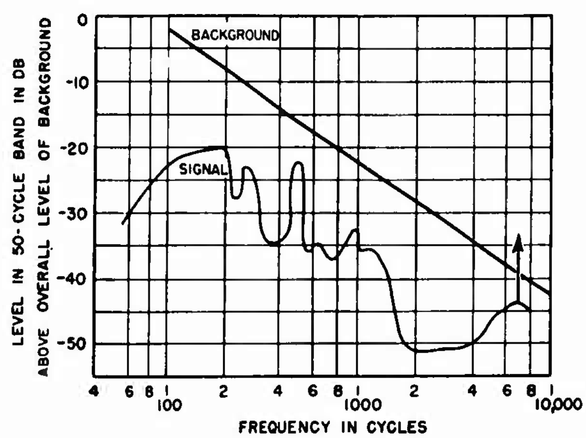


FIGURE 29. Audibility of sonic noise from a fleet-type submarine at 5 knots and periscope depth, masked by simulated deep-sea noise. *RD* for presentation band shown: —14 decibels. Character of primaudible signal: squeaking and rhythmic hiss. Primaudible frequency: 7 kilocycles. Arrow indicates fluctuating single-frequency component. See Table 4, line 7.

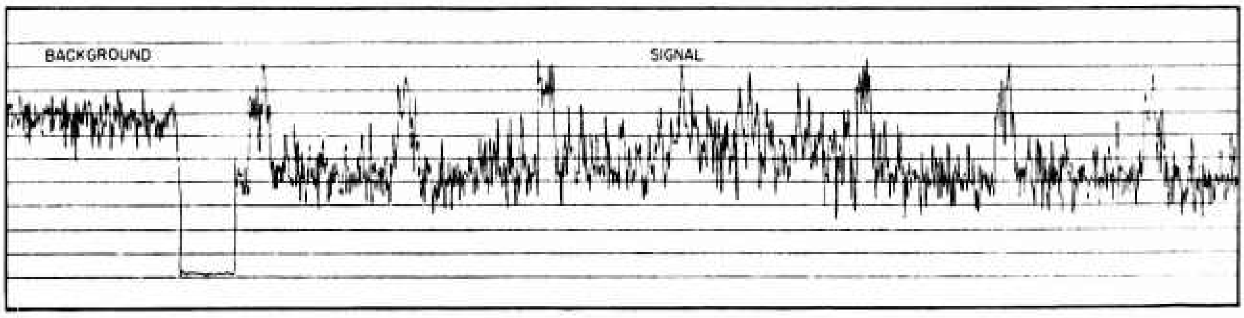


FIGURE 30. Recorder traces of components in optimal 50-cycle band centered at 7 kilocycles for sonic noise from a 5-knot submarine and simulated water noise (Figure 29). Vertical distance between lines represents power level change of 5 decibels. Total writing time was about 60 seconds. The relative level of signal and background corresponds to primaudibility.

nent tone at 7 kilocycles, there would be no substantial increase in the level of the signal at that frequency if the measurement had been made with a 400-cycle filter, which is the critical band width at 7 kilocycles, but the level of the distributed noise spectrum would have been increased by 9 decibels (10 log 400/50 = 9 decibels). Thus, the resultant signal-to-noise ratio, corrected to the appropriate critical band, is −1 decibel, rather than +8 decibels.

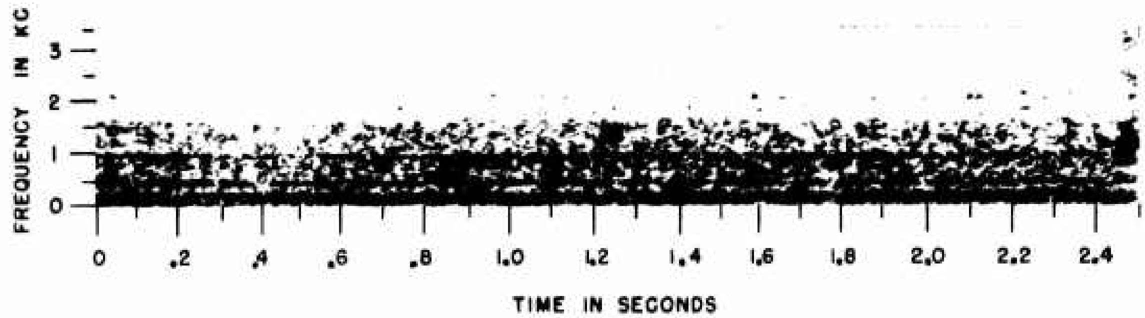


FIGURE 31. Time-frequency-intensity analysis for sonic noise from a 5-knot submarine (Figures 29 and 32).

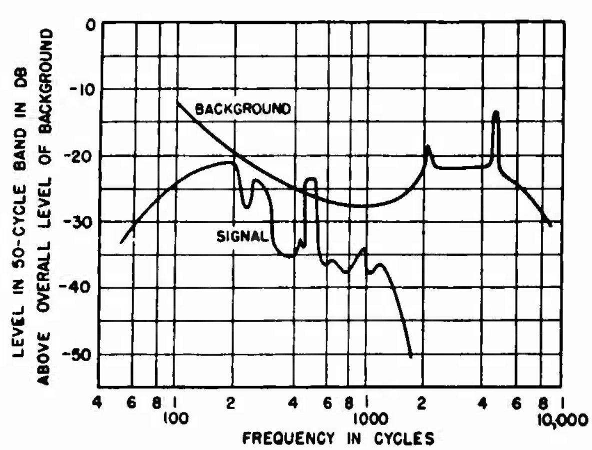


FIGURE 32. Audibility of sonic noise from a fleet-type submarine at 5 knots and periscope depth, masked by shrimp noise. *RD* for presentation band shown: —15 decibels. Character of primaudible signal: whine. Primaudible frequency: 500 cycles. See Table 4, line 22.

Figure 38 (line 13 in Table 4) represents a borderline case. The optimal frequency band for this signal was centered near 8.6 kilocycles. Recognized as a "swish," the primaudible component was definitely not a single frequency, but the signal spectrum was found to change markedly in level over the half-octave band between 6.4 and 9.0 kilocycles; in other words, signal-to-noise ratios in the 50-cycle and the half-octave bands did not agree. Furthermore, the signal-to-noise ratio of 1 decibel in the optimal 50-cycle band is too low if the primaudible component was narrower than 500 cycles, which is the critical band width at 8.6 kilocycles. To determine whether the critical-band criterion was met in this case, it would be necessary to use filters intermediate in width between 50 cycles and half an octave, and these were not available. Integration of power levels over the ten successive 50-cycle intervals in the signal which define the critical band centered at 8.6 kilocycles is not a practicable substitute for a single 500-cycle filter because the fraction of the energy passed by any of the 50-cycle filters amounts to 10^{-4} of the overall signal energy, whereas the latter fluctuates by at least 50 to 100 per cent during the course of a measurement. Thus, the experimental error is of the order of the effect which it is desired to study.

The results indicate that the recognition frequency at primaudibility is not a fixed characteristic of the signal but depends also upon the composition of the background. Thus, when the same signal is presented against two very different backgrounds, the optimal frequency band may shift from 0.2 to 8.6 kilocycles (compare Figures 38 and 40), or from 0.5 to 7.0 kilocycles (compare Figures 29 and 32). Similarly, the wide-band *RD* is not a fixed charac-

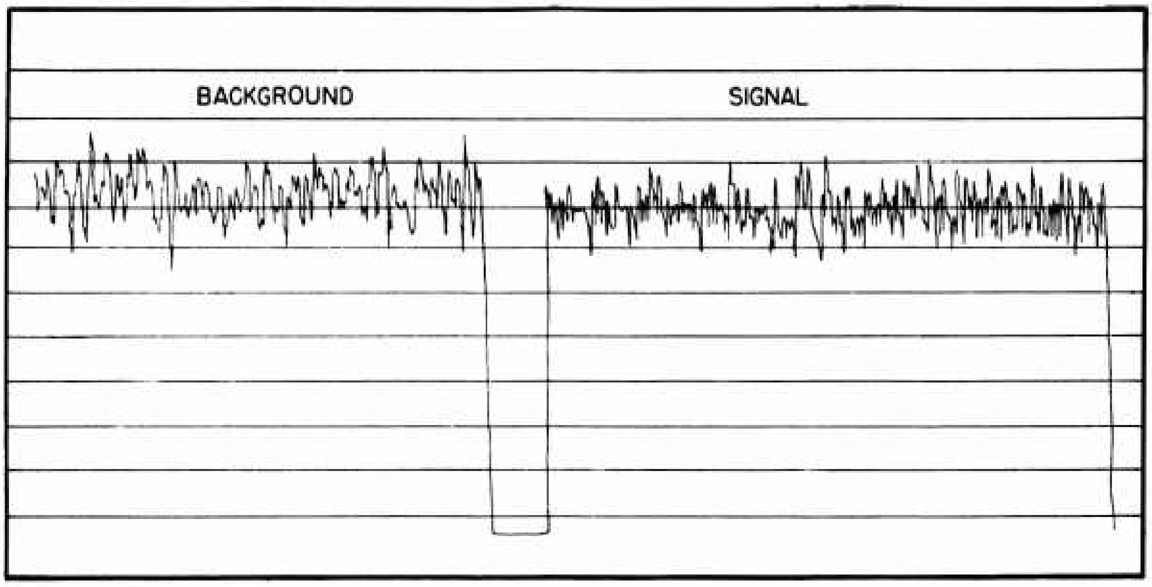


FIGURE 33. Recorder traces of components in optimal 50-cycle band centered at 500 cycles for sonic noise from a 5-knot submarine and shrimp noise (Figure 32). Vertical distance between lines represents power level change of 5 decibels. Total writing time was about 20 seconds. The relative level of signal and background corresponds to primaudibility.

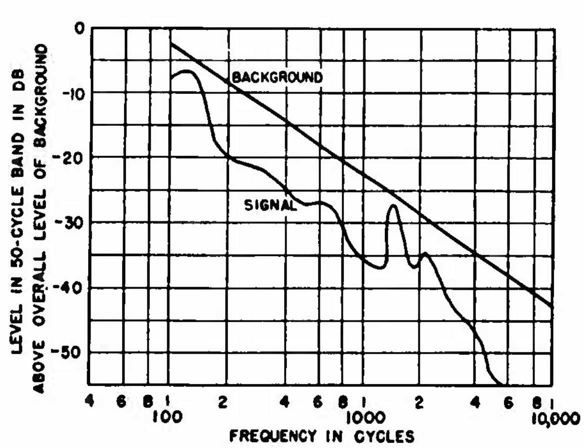


FIGURE 34. Audibility of sonic noise from a surfaced fleet-type submarine under way at 12 knots, masked by simulated deep-sea noise. *RD* for presentation band shown: —6 decibels. Character of primaudible signal: tinny grinding and low rumble. Primaudible frequencies: 120 and 1,400 cycles. See Table 4, line 9.

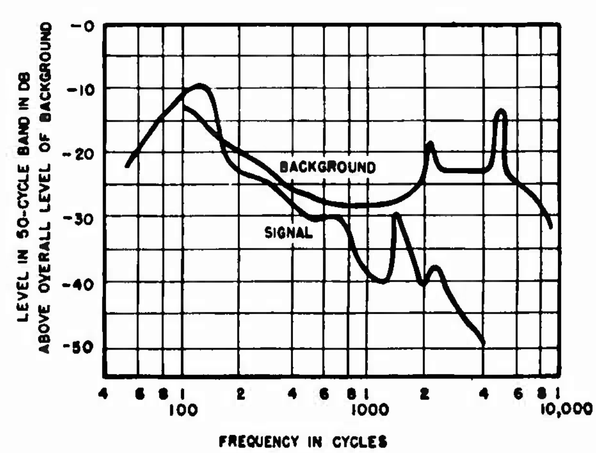


FIGURE 35. Audibility of sonic noise from a surfaced fleet-type submarine under way at 12 knots, masked by shrimp noise. *RD* for presentation band shown: —7 decibels. Character of primaudible signal: low rumble and whirr. Primaudible frequencies: 120 and 1,400 cycles. See Table 4, line 24.

teristic of signals and may vary more or less widely and erratically from one background type to another (compare, for example, lines 1 and 17 and also 2 and 18 in Table 4, to select two cases at random). On the other hand, the signal-to-noise ratio in the optimal critical band is much more nearly a constant quantity.

A few of the observations indicate that the recognition differential in the critical band may depend on background type to some extent. Thus, the same signals require 2 to 3 decibels more gain in the optimal critical band when the background is shrimp noise than when it is simulated ambient noise (compare Figure 34 with Figure 35 and also Figure 36 with Figure 37). Similarly, it was observed that the recognition differentials (measured at the filter input) were 1 to 2 decibels more favorable for various signals masked by shrimp noise when an 1,100-cycle low-pass filter was used, and that listening comfort was increased. While it has been suggested that some low-frequency masking results from intermodulation products (difference frequencies) of the high-frequency energy of shrimp crackle, it will be recalled that similar results were obtained with backgrounds containing much less high-frequency energy than shrimp noise (see Figure 10D), and therefore contributing much less intermodulation energy to the low-frequency region. It seems more likely that the help afforded by passing only the low-frequency components in such cases may be ascribed almost entirely to reduction of the loudness of the filtered sound, thereby permitting the optimal component to be presented at a higher listening level. This is especially important in cases where the optimal frequency region lies

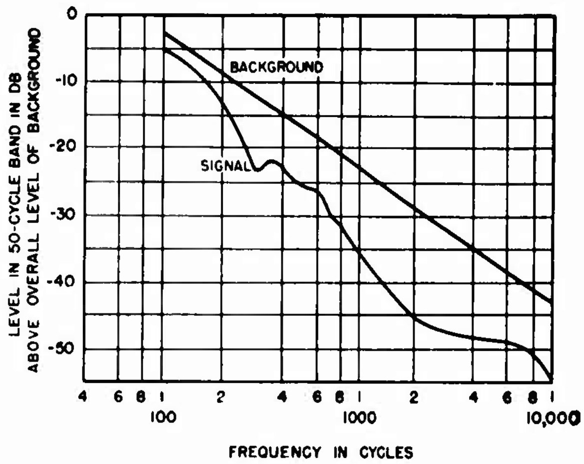


FIGURE 36. Audibility of sonic noise from an anchored S-class submarine, charging batteries with diesels, masked by simulated deep-sea noise. *RD* for presentation band shown: —6 decibels. Character of primaudible signal: low-frequency rhythm. Primaudible frequency: 120 cycles. See Table 4, line 1.

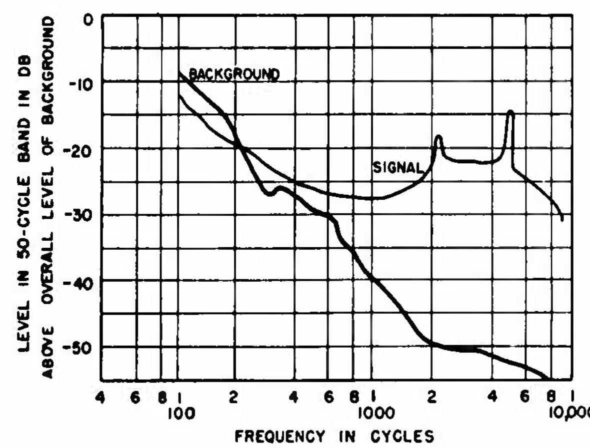


FIGURE 37. Audibility of sonic noise from an anchored S-class submarine, charging batteries with diesels, masked by shrimp noise. *RD* for presentation band shown: —10 decibels. Character of primaudible signal: low-frequency rhythm. Primaudible frequency: 150 cycles. See Table 4, line 17.

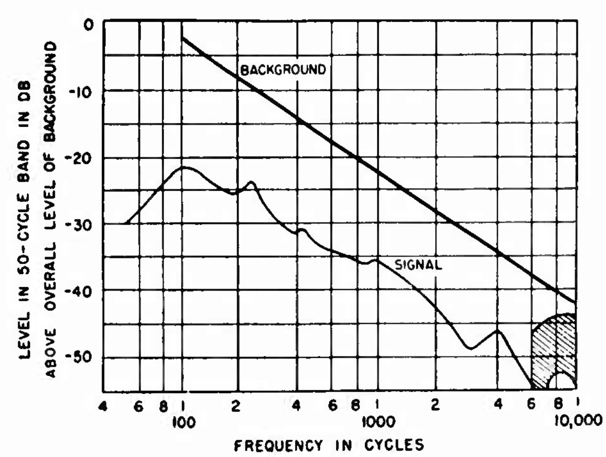


FIGURE 38. Audibility of sonic noise from a 7,000-ton freighter at 10 knots, masked by simulated deep-sea noise. *RD* for presentation band shown: —12 decibels. Character of primaudible signal: rhythmic swish. Primaudible frequency: 8.6 kilocycles. Shaded area indicates fluctuating components. See Table 4, line 13.

below 200 cycles, since acoustic leakage about the headphone cap, as well as smaller attenuation of room noise, tend to reduce listening efficiency. On the other hand, use of low-pass filters and increased listening level may fail to help when the optimal band lies at the lower frequencies, since headphones generally produce serious distortion in this region, resulting in a confusing similarity between signal and background.

While the sounds which were examined in the study under discussion indicate the usefulness of the critical-band concept, it should be observed that these sounds were, to some extent, arbitrary. Thus, the spectra shown in the figures represent electrical outputs from a large variety of hydrophones and amplifiers. Hence the wide-band recognition differentials obtained with signal and background pairs from unmatched hydrophones may give an unreliable picture of the field situation in which, of course, signal and noise are transmitted by the same system. For example, lines 17 and 18 in Table 4 show a wide band *RD* of −20 decibels for a submarine signal when the signal and background were recorded through widely different hydrophones, and an *RD* of −10 decibels when the same submarine signal was recorded through a hydrophone similar to that used in recording the noise. Clearly, the latter *RD* is the quantity of practical use. Unmatched recording hydrophones also account for the change in optimal component and improvement in *RD* shown by line 2 as compared with line 1 in Table 4.

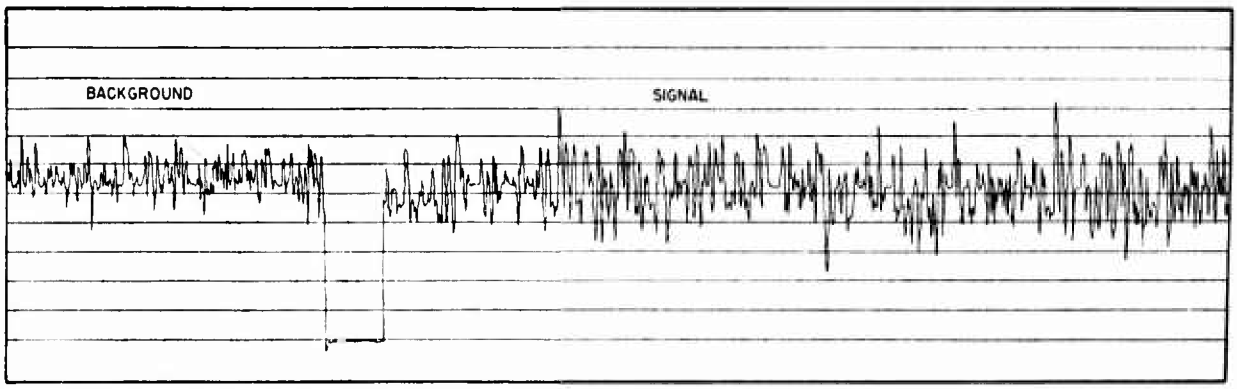


FIGURE 39. Recorder traces of components in optimal 50-cycle band centered at 8.6 kilocycles for sonic noise from a 10-knot freighter and simulated water noise (Figure 38). Vertical distance between lines represents power level change of 5 decibels. Total writing time was about 40 seconds. The relative level of signal and background corresponds to primaudibility.

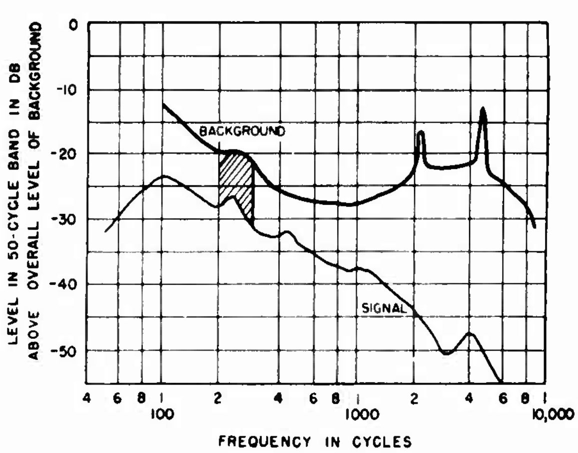


FIGURE 40. Audibility of sonic noise from a 7,000-ton freighter at 10 knots, masked by shrimp noise; shaded area indicates fluctuating components. *RD* for presentation band shown: —14 decibels. Character of primaudible signal: rhythmic chug. Primaudible frequency: 260 cycles. See Table 4, line 26.

Similarly, ground noise due to imperfections in the recording may introduce errors in relative level at the higher frequencies.[9,10] In addition, no examples of sonic self-noise are included in the group of masking backgrounds studied in the present case. The possibility that strong tonal components, generated by the operation of auxiliaries, may frequently occur in self-noise poses a question which cannot be answered on the basis of the present findings, since beats between tones in signal and noise probably render the simple critical-band concept inadequate (see, for example, Section 2.1.2). Furthermore, the diminution in signal level with increasing range is invariably accompanied, in practice, by selective attenuation, which weakens the higher frequencies to a greater extent than the low frequencies. The frequency composition of recorded signals used in tests of this kind, however, does not vary in a comparable way as a function of signal level. It seems quite likely, therefore, that in operational listening the prominent high frequencies in shrimp noise would have but little influence in masking sonic ship signals, since relatively little high-frequency energy is received from the latter at practical ranges. Finally, it is impossible to say that the spectra of the various sources employed in these tests are representative samples; in particular, the cases discussed down to this point give no adequate picture of the recognition of propeller sounds. The latter type of signal is discussed in the following section.

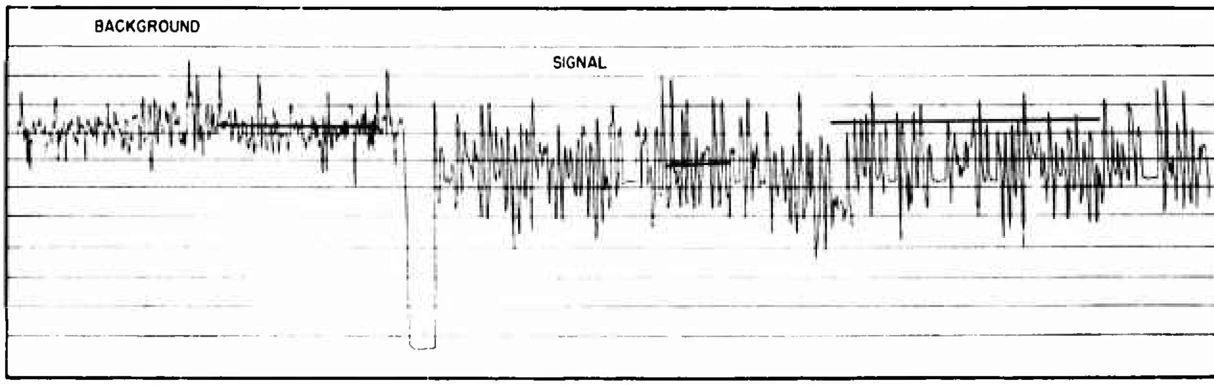


FIGURE 41. Recorder traces of components in optimal 50-cycle band centered at 260 cycles for sonic noise from a 10-knot freighter and shrimp noise (Figure 40). Vertical distance between lines represents power level change of 5 decibels. Total writing time was about 30 seconds. The relative level of signal and background corresponds to primaudibility.

4.2.3 Sonic Propeller Sounds

The information at present available concerning recognition of sonic propeller sounds is somewhat limited. Some quasi-field studies have beer made with these sounds and are described in reference 13. In the present section, some preliminary UCDWR tests, made partially with simulated signals, are described. The propeller simulations were amplitude-modulated thermal noise which had a spectrum slope of 0 decibel per octave. One of the masking backgrounds used was simulated deep-sea ambient, that is, unmodulated thermal noise

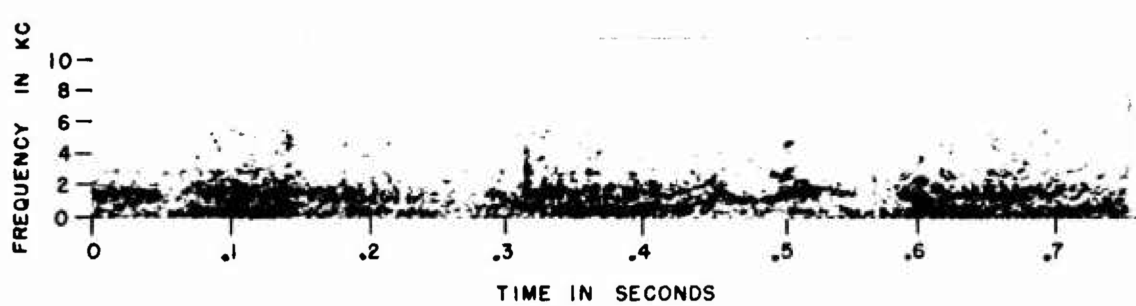


FIGURE 42. Time-frequency-intensity analysis for sonic noise from a 10-knot freighter (Figures 38 and 40).

with a slope of −6 decibels per octave. Both sounds are fairly good facsimiles of the actual ones and probably offer a useful guide to the factors affecting audibility of propeller cavitation. Tests were conducted in which both signal and noise extended over the entire 0.1- to 10-kilocycle band, and also in which filters were introduced to restrict the frequency range of the signal alone or of the entire mixture. The other masking background used was unmodulated wide-band thermal noise with a slope of 0 decibel per octave. Time patterns of these artificial signal and noise types are shown in Figures 62 and 63. It will be observed from these figures that the sole difference between background and signal was the modulation of the latter. Thus, the listening problem which is presented is that of distinguishing between the fluctuations normally present in the background and those caused by the presence of the signal. Since the time pattern of the background is random, any rhythm in the signal provides a useful cue.

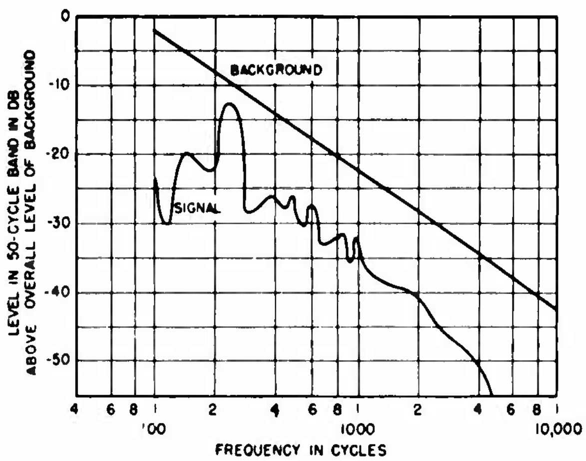


FIGURE 43. Audibility of sonic noise from a 2,870-ton destroyer, masked by simulated deep-sea noise. *RD* for presentation band shown: —11 decibels. Character of primaudible signal: low-pitched tone with propeller rhythm. Primaudible frequency: 250 cycles. See Table 4, line 14.

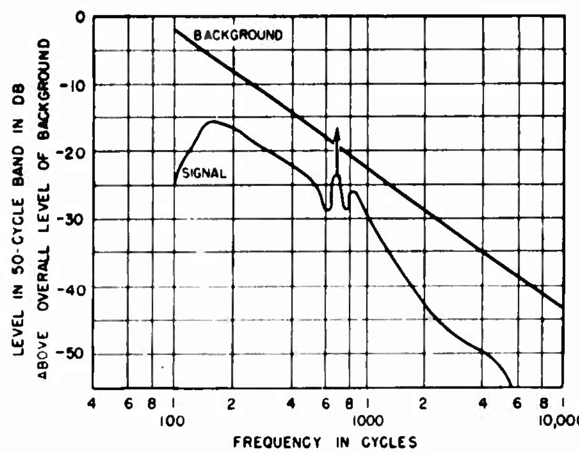


FIGURE 44. Audibility of sonic noise from tanker, masked by simulated deep-sea noise; the arrow indicates a fluctuating single-frequency component. *RD* for presentation band shown: —9 decibels. Character of primaudible signal: low rhythmic noise and whine. Primaudible frequency: 680 cycles. See Table 4, line 12.

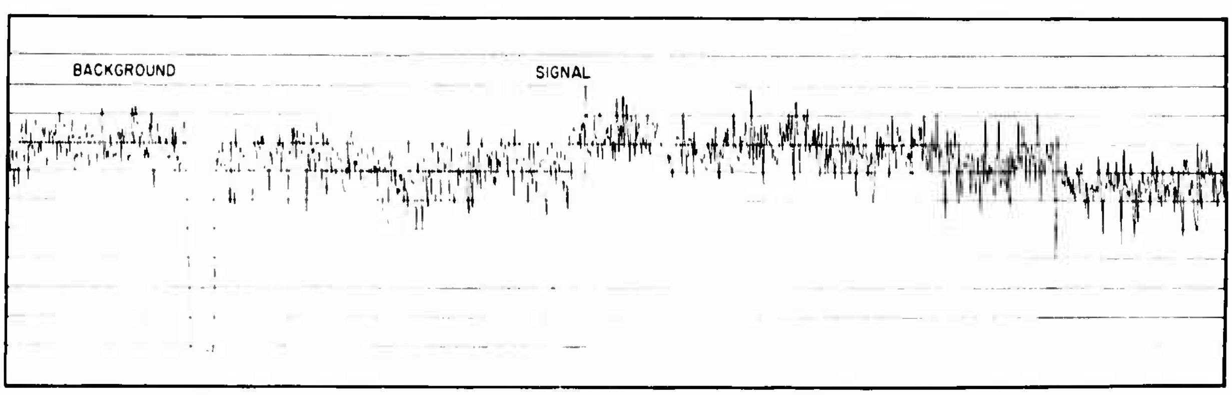


FIGURE 45. Recorder traces of components in optimal 50-cycle band centered at 680 cycles for sonic noise from a tanker and simulated water noise (Figure 44). Vertical distance between lines represents power level change of 5 decibels. Total writing time was about 30 seconds. The relative level of signal and background corresponds to primaudibility.

The general psychoacoustic data discussed in Chapter 2 provide a basis for understanding and evaluating these tests on the masking of propeller sounds. Previous observations (Section 2.2.2) indicate that rhythmic fluctuations can be readily perceived if the modulation rate does not exceed 10 cycles; at higher rates, the modulation is heard as a "flutter" and becomes more difficult to discern. For very low modulation rates, less than about 1 per second, a memory image of the intensity is required; since this image fades with time, recognition tends to deteriorate. It was also pointed out that, when both signal and noise are wide-band sounds with identical spectra, the recognition of the signal depends only on the general increase of intensity which it produces over the entire band. Under these conditions an intensity increment of about 1 decibel can be just recognized, at the optimal modulation rate of 3 cycles. Under ideal conditions very much smaller intensity increments can be perceived —even as small as 0.1 decibel. As already pointed out in Section 2.2.2, recognition of these very small increments is not likely in most practical situations. The data obtained in the masking studies described here give support to this view.

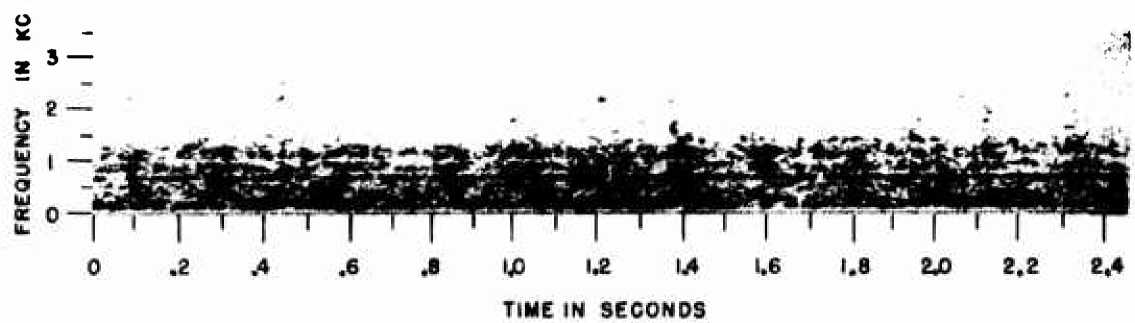


FIGURE 46. Time-frequency-intensity analysis for sonic noise from a tanker (Figure 44).

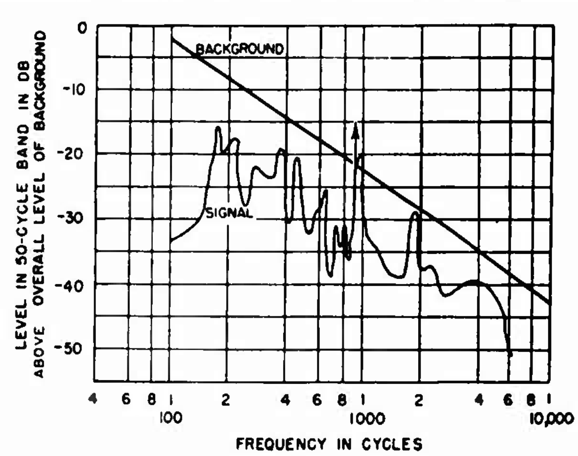


FIGURE 47. Audibility of sonic noise from a 45-knot torpedo (3-turbine type), masked by simulated deep-sea noise; the arrow indicates a fluctuating single-frequency component. *RD* for presentation band shown: —11 decibels. Character of primaudible signal: high-pitched whine or hum. Primaudible frequency: 970 cycles. See Table 4, line 15.

A minimal intensity increment of 1 decibel corresponds to a recognizable signal which is considerably weaker than the background. The recognition differential for such a signal may be estimated as follows. Let I be the average overall intensity of the background, and let i_p and i_m be the peak and minimum intensities of the signal. If m is the percentage amplitude modulation of the signal, then, since intensity is proportional to the square of the amplitude,

$$\frac{i_p}{i_m} = \left(\frac{1+m}{1-m}\right)^2. \tag{9}$$

Suppose recognition occurs when the maximum intensity $I + i_p$ of the signal-background mixture is 1 decibel above the minimum intensity, $I + i_m$. Since 10 log 1.25 equals 1 decibel, the condition for detecting the change is

$$1.25 = \frac{1 + i_p}{1 + i_m}. \tag{10}$$

If i_m is eliminated from these two expressions, we obtain after some manipulation

$$\frac{I}{i_p} = \frac{20m}{(1+m)^2} - 1 \quad . \tag{11}$$

Also, the mean intensity i is simply $(i_p + i_m)/2$, which, together with equation (9), gives

$$i = i_p \frac{1 + m^2}{(1 + m)^2}. \tag{12}$$

Thus, for 100 per cent modulation ($m=1$), I is four times i_p, giving a peak intensity 6 decibels below the average noise background. Since the mean signal intensity is 3 decibels below the peak for 100 per cent modulation (see equation (12)), the RD based on average signal level is, in this optimal case, -9 decibels. The latter conclusion is not obvious from a power level record of the signal since the minimum intensity of a sound with 100 per cent amplitude modulation is zero, or an infinite number of decibels below reference, as indicated by the dotted portion of the curve for $m=1$ which is given in Figure 54A. For 50 per cent modulation, the peak is 9/31 of the average noise, or 5.4 decibels below; the corresponding RD based on average signal intensity is -7.9 decibels. For 30 per cent modulation the peak signal level is 4.1 decibels below the noise, giving a mean level RD of -6.0 decibels. These relationships are portrayed in Figure 54, where the signal is shown in relation to the noise background. At the top of each diagram the combination of noise and modulated signal is drawn, with a total variation of 1 db as assumed.

Three limitations are apparent in the above analysis. In the first place, the signal and background must be wide-band noises, with random phases. If both signal and background were pure tones, for example, the intensity of the two would be obtained by adding their amplitudes and squaring rather than adding intensities; the RD in this hypothetical case may be shown to be -24 decibels (see Section 2.1.2). Since, in fact, the signal and background are very unlikely to have common phases at any frequency, the results shown in Figure 54 should be moderately realistic. In the second place, the above result does not hold for weak modulations—less than about 10 or 15 per cent.

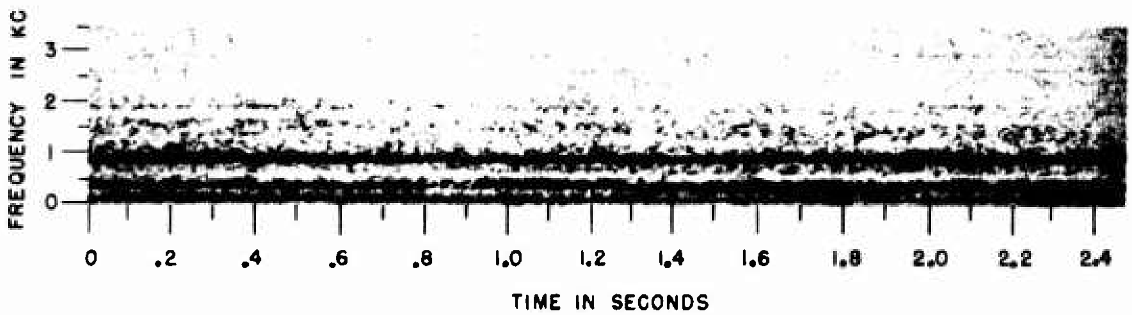
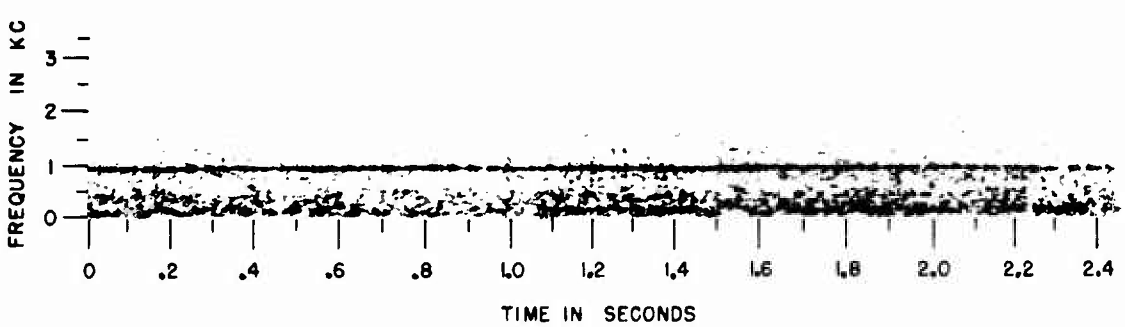


FIGURE 48. Time-frequency-intensity analysis for the signal and the signal-background mixture shown in Figure 47. For the signal-background mixture (lower trace), relative intensities have been adjusted to primaudibility.



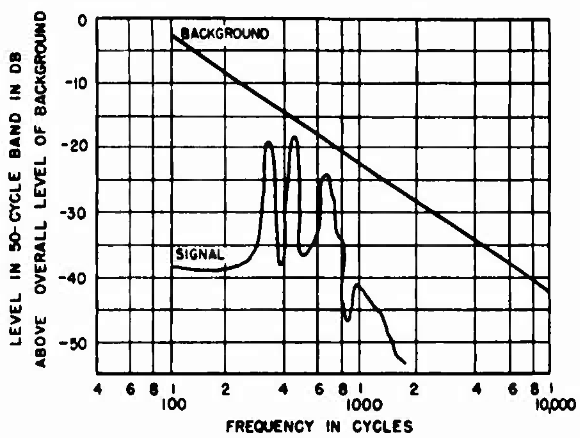


FIGURE 49. Audibility of sonic noise from a 30-knot torpedo (electric type), masked by simulated deep-sea noise. *RD* for presentation band shown: —17 decibels. Character of primaudible signal: hum and whine. Primaudible frequency: 660 cycles. See Table 4, line 16.

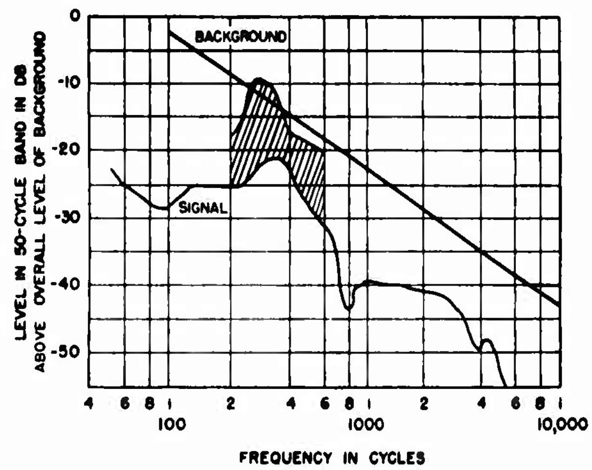


FIGURE 50. Audibility of sonic noise from a German Type VIIC submarine (GRAF) at 2½ knots and periscope depth, masked by simulated deep-sea noise; the shaded area indicates fluctuating components. *RD* for presentation band shown: —11 decibels. Character of primaudible signal: low rhythmic bubbling. Primaudible frequency: 300 cycles. See Table 4, line 10.

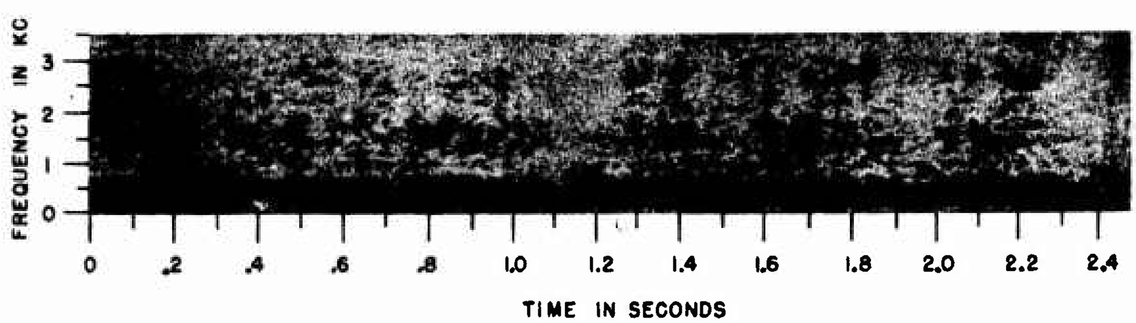


FIGURE 51. Time-frequency-intensity analysis for sonic noise from a 2.5-knot submarine (Figure 50).

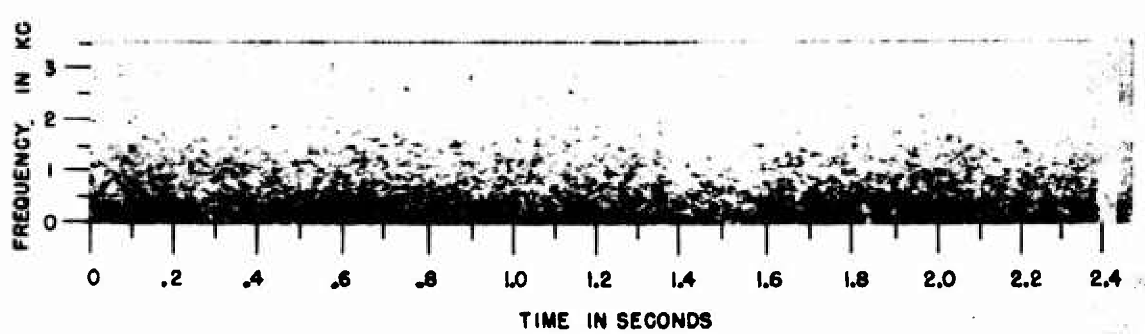


FIGURE 52. Time-frequency-intensity analysis for the signal-background mixture shown in Figure 50, with relative intensities adjusted to primaudibility.

Such a small variation would be difficult to detect, and the signal would be detected by sweeping on and off target rather than by listening for characteristic modulation at each bearing. Finally, both signal and background must have parallel spectra over a wide range of

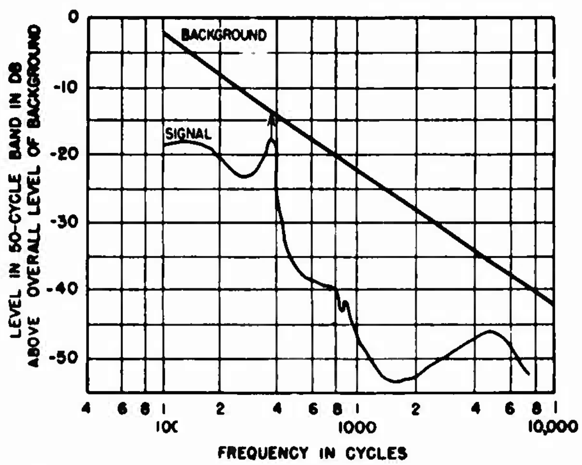


FIGURE 53. Audibility of sonic noise from a convoy (13 merchant vessels, 6 large tank lighters, and 3 PC escorts), masked by simulated deep-sea noise; the arrow indicates a fluctuating single-frequency component. *RD* for presentation band shown: —12 decibels. Character of primaudible signal: hum or whine. Primaudible frequency: 390 cycles. See Table 4, line 11.

frequencies with no tonal components in the signal, so that recognition occurs in a wide band.

Propeller sounds are usually composed of wide-band cavitation noises with a spectrum similar to the background over a fairly wide band, usually several kilocycles or more. If attention is confined to this band, the results derived may be expected to be applicable provided that the modulation is 30 per cent or more. Thus the spectrum differential, or the spacing between the signal and noise spectra at primaudibility, for such modulated sounds should be between -5 and -9 decibels, provided that the modulation rate lies between 1 and 10 cycles (see Figure 64) and provided that tonal components (or other peaks) in different parts of the signal spectrum are not the primaudible components. Also, the background noise level must be reasonably steady for this result to hold. If the time pattern of the background noise is similar to that of the signal, as may perhaps occur with certain types of self-noise, a spectrum differential more nearly equal to zero would be expected. More usually, however, the background is not modulated and a spectrum differential between -5 and -9 decibels may be expected.

The available observations support this analysis in a number of situations. It has already been noted in Section 4.1.3 that the British data are in agreement with these results. When modulated propeller sounds could just be heard in the presence of wide-band masking noise, the smallest distance between the two spectra was 6 to 10 decibels. While signal and background spectra did not have identical shapes, they were sufficiently parallel above 2 kilocycles for the foregoing analysis to be relevant, and thus the agreement can be taken as confirmation of the theory. As already indicated in the discussion of Figure 54, the arithmetic mean of two intensities a and b is not generally equal to the intensity c which is obtained by averaging the intensity levels 10 log a and 10 log b. The latter process gives $(10 \log a + 10 \log b)/2 = 10 \log \sqrt{ab}$. In other words, the derived quantity corresponds to the geometric rather than the arithmetic mean. Consequently, use of the mean signal level, read from a power level trace, may yield somewhat lower values for the spectrum differentials than estimated above.

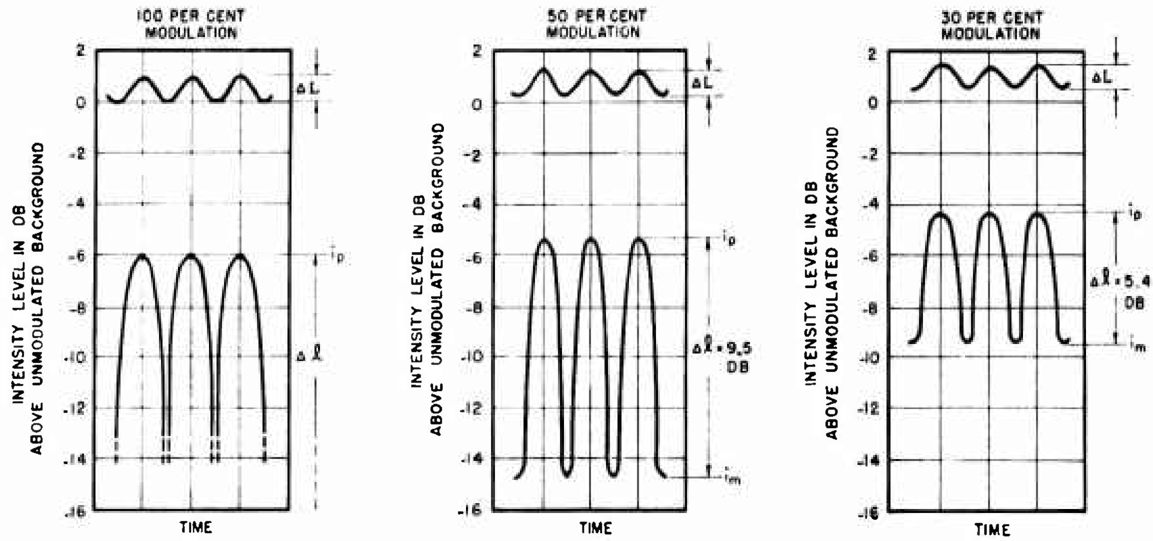
The UCDWR sonic data also show agreement with expectation. Thus, it was found that a recurrent change in the intensity of a wide-band sound can be detected when the ratio of maximum to minimum intensity equals 1.25, or 1 decibel, and that this change can be detected equally well when it is produced either by modulating a single thermal noise source or by adding a modulated thermal noise to an unmodulated one having identical band width and spectrum slope.

Similar results were obtained with ship sound recordings. Figures 55 and 56, for example, illustrate the relation at primaudibility between the spectra of a submerged submarine and of simulated deep-sea ambient. In this case, the primaudible band included the frequency region between 2 and 7 kilocycles. Within this region, the average signal level was 12 decibels below average noise, and the maximum signal level 6 decibels below noise. The primaudible signal perceptibly raised the total loudness and was heard as a rhythmic thrash with a clearly audible modulation of about 2 beats per second.

It was found, when the components of the signal-background mixture with frequencies between 2 and 7 kilocycles were transmitted through progressively narrower band-pass filters, that the primaudible signal level rose to continuously higher values. With a 50-cycle band-pass filter centered near 5 kilocycles, the signal was primaudible when its average level was equal to the average noise level, in other words, when the maximum signal level exceeded the noise. When listening through the 50-cycle filter, the amount of random fluctuation perceived by the ear was noticeably greater than when listening in the wider bands. In addition, the mixture had an unpleasant tonal quality when passed by the narrow filter, the amplitude modulation was difficult to discern, and no reliable turn count could be obtained at primaudibility.

Figure 57 sheds some light on the preceding observations. The three rows of recorder traces show time patterns of simulated propeller noise, flat thermal noise background, and a mixture of both. The latter was presented to the recorder at the same signal-to-noise ratio for all the admittance band widths illustrated. The particular signal-to-noise ratio used is several decibels higher than is needed for primaudibility when the mixture is heard through the 50-cycle filter. The input levels of signal and noise were held constant, and, since filtering reduces the energy fed to the recorder, the gains were adjusted to keep the traces in the same relative positions on the paper. Such gain changes do not affect the relative amount of fluctuation shown for a particular sound.

The observed fluctuation does increase with decreasing band width. This observation may account, at least in part, for the impaired audibility of the submarine cavitation which was just discussed, and it indicates that restrictions of system band width must be made with caution. This same problem is touched upon in



$\triangle L$ = 10 log $[(I + i_p) / (I + i_m)]$ = recurrent variation in level of mixture (1 db)
$\triangle l$ = 10 log (i_p / i_m) = recurrent variation in level of signal
I = Mean intensity of unmodulated background (reference intensity)
i = Mean intensity of modulated signal
m = Percent amplitude modulation of signal
i_p = Recurrent peak intensity of modulated signal = $i(1 + m)^2$, since intensity is proportional to square of amplitude
i_m = Recurrent minimum intensity of modulated signal = $i(1 - m)^2$

FIGURE 54. Relative intensity levels of a sinusoidally modulated signal and of the signal-background mixture. Lower curves represent time patterns of signals; upper curves represent time patterns of mixtures.

various other connections; see, for example, the discussion in Chapter 8. It may be mentioned here, however, that a sharply tuned system has a longer response time than one which is untuned or broadly tuned, and the noise fluctuations in such a tuned system are less rapid than in a system passing a broad band of frequencies. In other words, admittance band width determines the "resolution time" of the system.

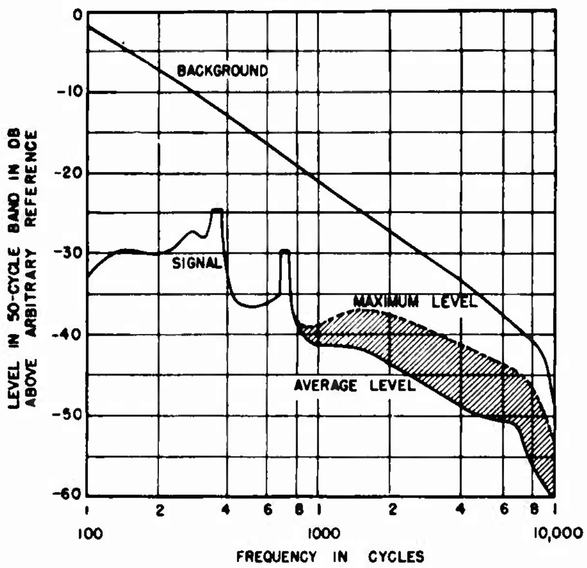


FIGURE 55. Audibility of sonic noise from a fleet-type submarine at 120 rpm (6 knots) and periscope depth, masked by simulated deep-sea noise. Character of primaudible signal: rhythmic thrash, twice each second.

The recorder traces under discussion show the influence of two distinct resolution times: that of the electrical system and that of the recorder. The latter instrument indicates the average d-c voltage applied to its terminals during its own response time. The greater "smoothness" of the trace obtained with reduced electrical selectivity means that the broader systems feed a greater number of fluctuations to the recorder during its resolution time and that the larger sample shows less mean deviation from the average noise level. To a first approximation, the degree of fluctuation should be inversely proportional to the square root of the band width.

One further point is worth noting. It is evident that if the critical bands were completely independent, narrow filtering should not have impaired recognition of the submarine cavitation. The signals received in each "critical-band filter" of the ear are therefore not heard sepa-

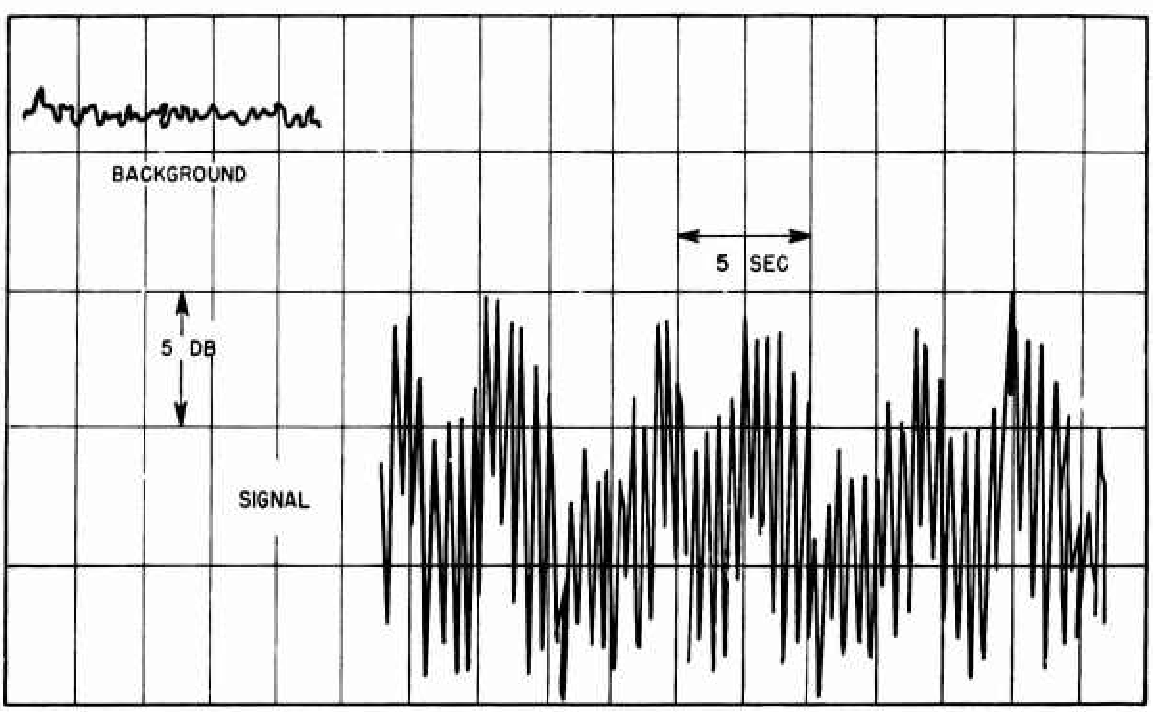


FIGURE 56. Recorder traces of components in optimal band for the sonic noise from a 6-knot submarine and the simulated deep-sea noise shown in Figure 55. The relative level of signal and background corresponds to primaudibility.

rately, but as a unit. The ear's response to the mixture admitted by the 9-kilocycle band seems to involve some fusion of the impulses received by various parts of the basilar membrane and tance of about 120 yards (see Figure 9 in Chapter 3). The masking background was simulated deep-sea noise, and recognition occurred when signal and background levels were

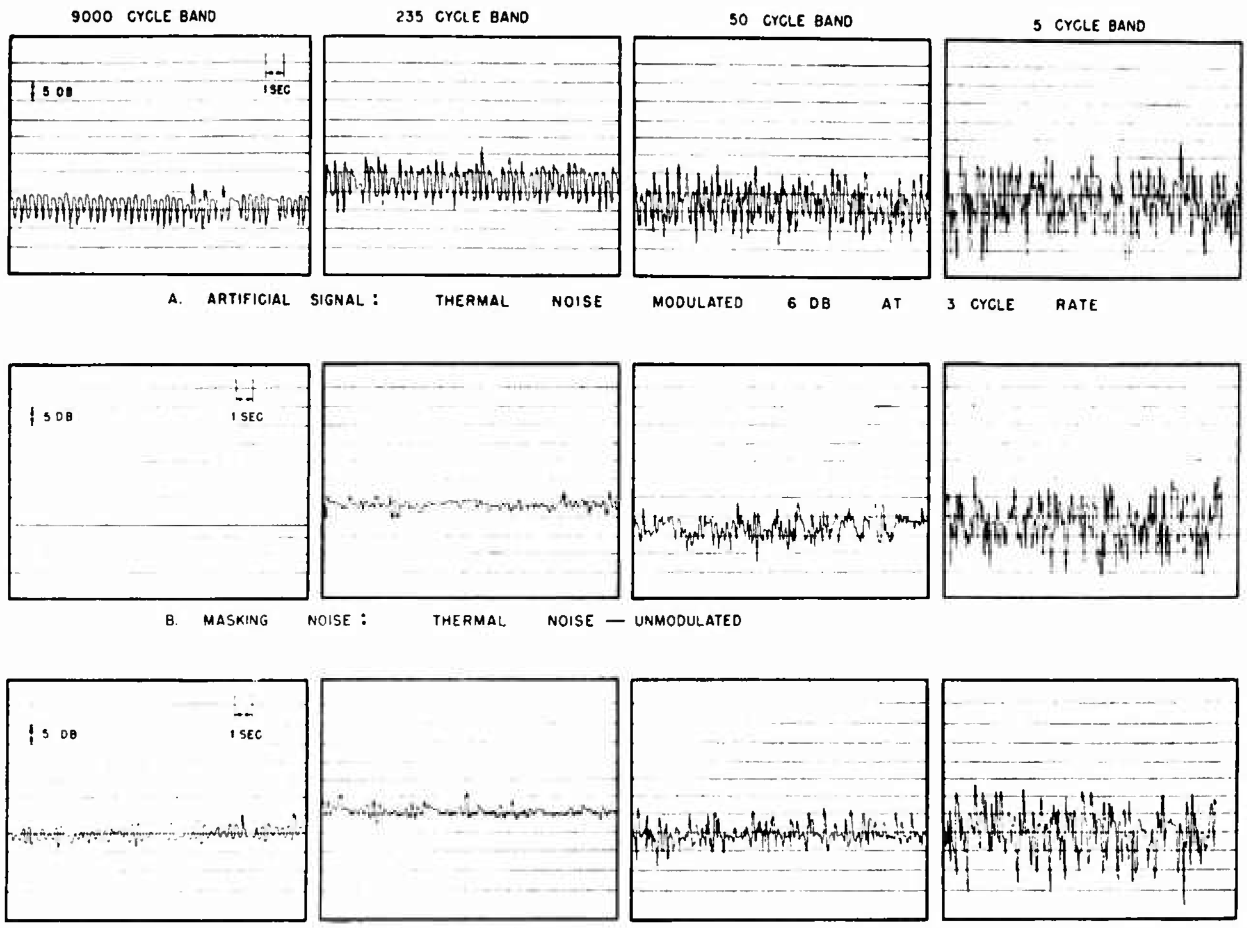


FIGURE 57. Power level recorder traces showing time-amplitude patterns of thermal noise.

is apparently more like the recorder trace for a 9-kilocycle band than it is like the trace for a band 50 or 235 cycles wide, although the latter bands are approximately the width of the aural critical bands. This integrating property of the ear is also apparent in the observed masking of FM pulses (see Section 8.1.3).

The foregoing discussion applies to masking of propeller noise when this consists of modulated cavitation sounds. Propeller noise is apparently not always of this character, however, as evidenced by the following two examples. Figure 58 shows the signal-background relationship at primaudibility for an aircraft carrier moving at 15 knots and recorded at a distance of about 120 yards (see Figure 9 in Chapter 3). The masking background was simulated deep-sea noise, and recognition occurred when signal and background levels were approximately equal at all frequencies in the entire listening band. This case stands at the opposite extreme from Figure 55. Such variability in the recognition of propeller sounds is perhaps the result of intrinsic variability associated with the presence or absence of strong cavitation. Thus the submarine whose spectrum is shown in Figure 55 was proceeding at 6 knots, which is well above the cavitation threshold at periscope depth, while the carrier was moving at a speed which produces no marked cavitation in her class of vessels.[11] It should be observed, however, that the listening test described by Figure 58 was performed without benefit of the sweep pattern which is introduced by training a directional hydrophone across the target bearing. With optimal hydrophone sweep rates (see Section 5.1), it should be possible to detect a source like the carrier

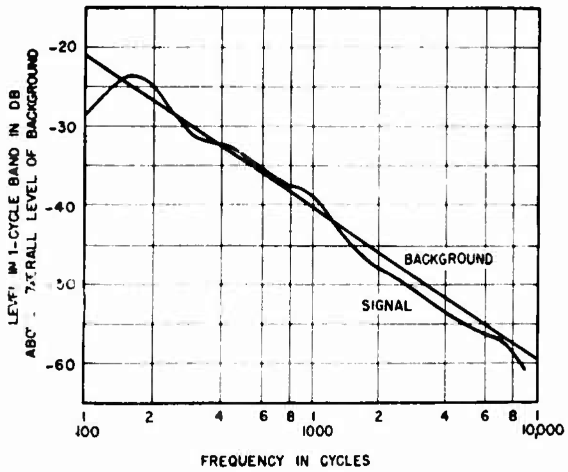


FIGURE 58. Audibility of sonic noise from 15-knot aircraft carrier *A*, masked by simulated deep-sea noise. *RD* for presentation band shown: —3 decibels. Character of primaudible signal: propeller sounds.

when the intensity increment produced in the transit is about 1 decibel, instead of the 3 decibels indicated by Figure 58, and, hence, when the signal is about 6 decibels below average noise. For signals like that generated by the submarine, sweeping the target bearing should produce little additional improvement; indeed, it is likely that hydrophone sweep rates should be diminished for well-modulated signals in order to obtain optimal results.

The spectrum of another carrier of the same class, operating under similar conditions, is shown in Figure 3 of Chapter 3. The dominant tone at 1,100 cycles was produced by a "singing" propeller, and was amplitude modulated at the screw rate. The "cavitation peak" below 300 cycles also appeared in the spectrum of the output of the first carrier when measured at bow and quarter aspects (see Figure 9 of Chapter 3); this trend has been ascribed to source directivity.

In the presence of simulated ambient noise, the second carrier was detected as a faint whine. The relation between signal and noise spectra at primaudibility, as measured in 50-cycle bands, is shown in Figure 59. The critical-band criterion predicts that a tone at 1.1 kilocycles will be primaudible when its level in a 50-cycle band is just equal to the distributed noise in this band. The apparent discrepancy is due to the fact that the spectrum, as plotted, gives the average, rather than the maximum level of the tone. It is evident that the second carrier is much more vulnerable to detection than the first, and may be better camouflaged acoustically through a simple modification of propeller design.

The carrier spectrum shown in Figure 59 has been recomputed to show how it would appear

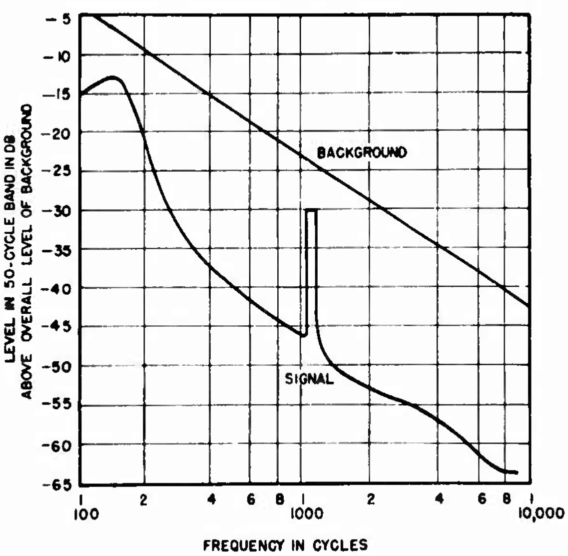


FIGURE 59. Audibility of sonic noise from 15-knot aircraft carrier *B*, masked by simulated deep-sea noise. *RD* for presentation band shown: —12 decibels. Character of primaudible signal: whine.

if the analysis had been made with octave filters exclusively. With such an analysis, the relation between signal and background spectra at primaudibility would be given by Figure 60. Comparison of the latter with Figures 6 through 23 indicates the general similarity of the British and the San Diego results, when allowance is made for differences in technique of analysis. It also demonstrates, incidentally, that wide-band analysis is a rather blunt tool for some purposes.

The observations discussed so far in this section apply exclusively to the situation in which signal and noise spectra have closely similar slopes over a wide frequency band. Since attenuation modifies the 6 decibels per octave slope of the propeller cavitation spectrum and since the self-noise received in ship-mounted hydrophones usually has a slope of between 9

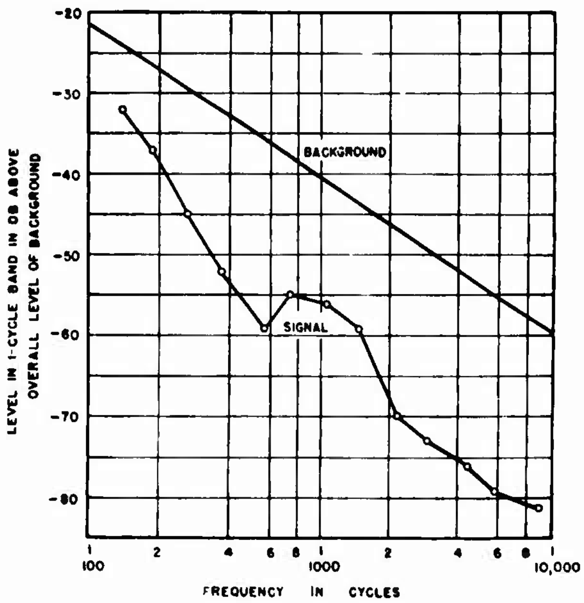


FIGURE 60. Spectra for the sonic noise from aircraft carrier *B* and simulated deep-sea noise shown in Figure 59, when the energy is measured in octave bands.

and 12 decibels per octave, it will not in general be true that signal and noise spectra are parallel over a wide frequency band. It is useful, therefore, to examine some further observations made with wide-band sounds generated in the laboratory.

The masking background in these tests was simulated deep-sea ambient, which had a slope of −6 decibels per octave and extended from 0.1 to 10 kilocycles (see Figure 62). The signals were 100 per cent amplitude modulated at a rate of 3 cycles (see Figure 63) and were obtained from a wide-band source of thermal noise with a flat spectrum (amplitudes equal at all frequencies). Each test was made with a different band-pass filter in the signal channel. These filters varied from half an octave to several octaves in width, and their midfrequencies fell at representative points in the interval between 500 and 7,000 cycles.

The signal-background relationship at primaudibility which was typically observed in these tests is shown in Figure 61. The primaudible group of signal components was invariably centered at the high-frequency limit of the signal pass band; and the maximum signal-to-average noise ratio in this optimal band was very nearly −2 decibels in all the cases studied. This signal-to-noise ratio is essentially the same as that found for fluctuating tonal components. In other words, the critical-band criterion applies also to the situation in which the primaudible sound consists of a narrow band of frequencies, and the ability to discriminate small differences of loudness is less than for the case in which the primaudible component is heard in a wide band (see Figure 55). The optimal signal-to-noise ratios observed in other masking tests in the series illustrated in Figure 61 showed a slight dependence on the

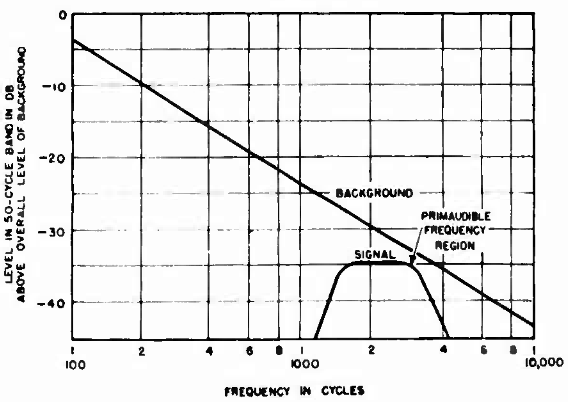


FIGURE 61. Audibility of modulated thermal noise, passed by an octave filter, masked by simulated deep-sea noise. *RD* for presentation band shown: —18 decibels.

frequency of the primaudible component and possibly also on sensation level. Since this set of tests was of a preliminary nature, the results are not precise enough to indicate whether the slight differences in optimal signal-to-noise ratio are more closely related to the critical band width function or to the loudness increment function. In any case, the observations discussed in the present section indicate that the width of the primaudible band affects the audibility of signals.

The rate of modulation may also be expected to affect the *RD* for a wide-band or narrow-band sound. No systematic studies of the role played by modulations found in underwater sounds have as yet been made. In lieu of this, however, a rough estimate of what may be expected in practice for wide-band sounds may be obtained from the data in Figure 16 of Chapter 2. This estimate, computed for a signal and a noise background with similar spectra, heard level of 25 decibels were used, since the total loudness of wide-band signals tends to limit the possible gain available in the optimal band. Furthermore, the data presented in Chapter 2

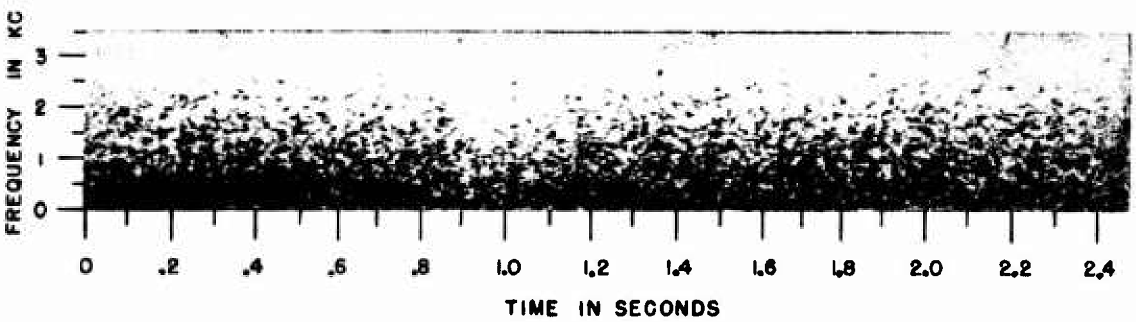


FIGURE 62. Time-frequency-intensity analysis for simulated water noise.

in a wide frequency band, is given in Figure 64.

The method of calculation has already been explained in connection with Figure 54. The ordinate, labeled "spectrum differential," shows the expected spacing, when the signal is just detectable, between the essentially parallel portions of the signal and noise spectra which lie in the primaudible frequency band. The abscissa gives propeller rpm, and the curve applies to 50 per cent amplitude modulation.

This curve has been drawn for the situation in which both signal and noise spectra are expressed in terms of average values of rms level because that has been the usual practice in measuring underwater sounds up to the present. Several assumptions have been made in preparing Figure 64. To begin with, the data in Figure 16 of Chapter 2 pertaining to a sensation level of 25 decibels were used... imply that the loudness increments selected for the present estimate are large enough to have a detection probability of 100 per cent rather than of 50 per cent, and even smaller changes can be detected when the sound is heard with both ears, as in most practical listening. As already pointed out, the more conservative values have been chosen, since the intensity limen (smallest perceptible intensity increment) varies with the primaudible frequency and because various other unfavorable factors may enter the field situation. In addition, the values of the intensity limen employed here were derived from studies with tones. General considerations, however, as well as a limited amount of experimental evidence (see Section 2.2.2) indicate that limens for tones and for distributed sounds are comparable in magni-

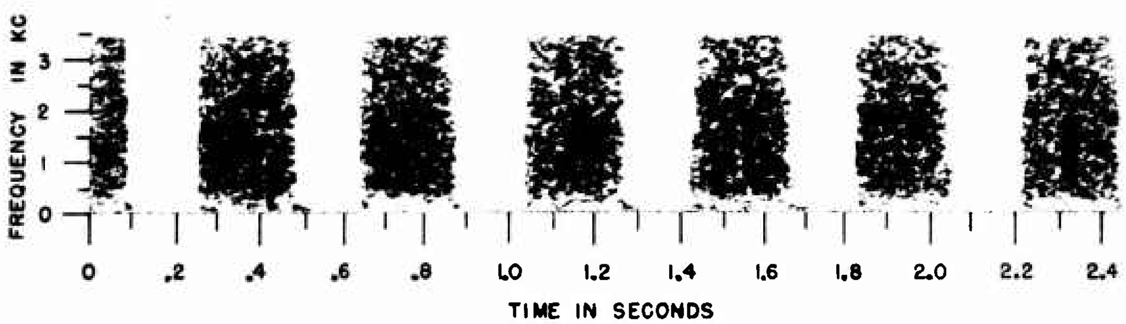


FIGURE 63. Time-frequency-intensity analysis for 100 per cent amplitude-modulated thermal noise.

tude, although the facts have not been established in detail. Finally, a modulation value of 50 per cent has been adopted as typical of well-modulated cavitation; for comparison, it should

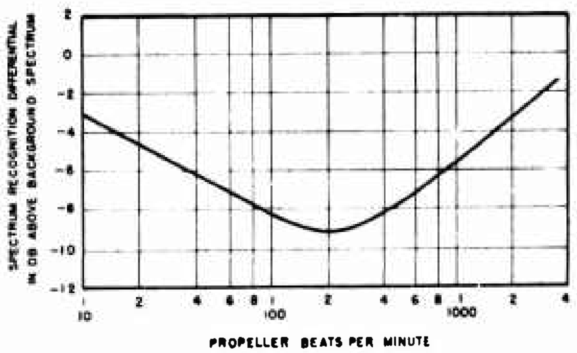


FIGURE 64. Estimated effect of the rate of amplitude modulation on the detectability of propeller sounds. This estimate assumes 50 per cent amplitude modulation and a wide band of primaudible frequencies.

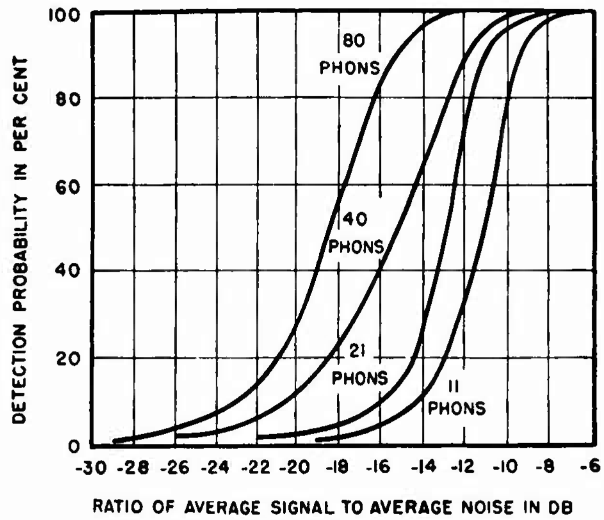
be noted that the curve in Figure 64 would be shifted up—to smaller negative values—by about 2 decibels for the case of 30 per cent modulation and would be shifted down by about 1 decibel for 100 per cent modulation.

Several practical points should be noted in connection with Figure 64. It applies primarily to recognition of propeller cavitation sounds and has no meaning below the cavitation threshold. The latter may occur at widely different propeller rates, depending on circumstances. In fact, it seems probable that the degree as well as the rate of modulation is a function of propeller rpm. While the indicated range of differentials checks fairly well with the available observations, it seems likely that the curve relating spectrum differential to screw rate for a particular vessel will, in general, follow a somewhat different course from that shown in Figure 64, because of the influence of at least four other factors.

In the first place, there is evidence[11] that the slope of the cavitation spectrum is a function of ship speed. Secondly, some propeller sounds seem to be characterized to a greater or lesser degree by the presence of frequency modulation; this is heard as a periodic swish,[g] and, in an analysis, shows up as a recurrent change in amplitude of the components admitted by each of a group of narrow filters tuned to neighboring frequencies (see Figure 12 in Chapter 3). The effect upon audibility of such a shifting spectral prominence depends on factors which have not yet been investigated.

The third factor referred to above is background type; the breaking of surf and similar sounds showing fairly long and regular recurrence times may counteract the effects of amplitude modulation in the signal. In that event, sweeping a directional hydrophone through the target bearing may help, although this device has a number of limitations (see Chapter 5). In general, hydrophone sweep should help most when the propeller modulation rate is either well below or well above the optimal rate of 180 rpm (see Figure 64).

The fourth factor is the loudness level at which the signal and background are heard. The estimated effect of loudness level (when



LEVEL IN PHONS	80 PER CENT RECOGNITION PROBABILITY	50 PER CENT RECOGNITION PROBABILITY	DIFFERENCE IN DB	n
80	-16.3	-18.5	2.2	2.7
40	-12.8	-15.2	2.4	2.5
21	-11.5	-13.0	1.5	4.0
11	-11.1	-9.9	1.2	5.0

FIGURE 65. Estimated effect of loudness level on the probability of detecting propeller sounds with 50 per cent amplitude modulation in the presence of a steady background.

[g] While no systematic observations are available, preliminary results obtained at UCDWR indicate that the ear does not readily distinguish between frequency and amplitude modulations in propeller sounds.

signal and background have similar frequency compositions) is shown in Figure 65, which has been recomputed from reference 12 by assuming that the intensity increments indicated in the latter were produced by the introduction of a signal, that is, by the method already described in the discussion of Figure 54. The very low recognition differentials shown in the figure correspond to the very small intensity increments which can be perceived under ideal conditions. Under practical conditions the 50 per cent *RD* will presumably correspond more nearly to a 1-decibel intensity increment. It may be noted that, when cavitation sounds are heard in only part of the spectrum, it is the contribution of these spectral components to the total loudness that is significant. Intense components at low frequencies, for example, may add to the total loudness without modifying the effective loudness of the cavitation sounds.

Two effects are shown in Figure 65: (1) by increasing the listening level from 11 to 80 decibels, the computed 50 per cent recognition differentials are improved by 9 decibels or more, in other words, smaller changes of intensity can be detected at the higher levels; (2) the computed transition curves are not only shifted but their slopes (see Section 4.1.4) also change more or less progressively, so that discrimination deteriorates more rapidly (with diminishing signal-to-noise ratio) at the low gain settings than at the high. The magnitudes of these changes are approximately independent of the assumed degree of signal modulation; thus, each curve is shifted to the left by about 1 decibel for the case of 100 per cent modulation and to the right by about 2 decibels for the case of 30 per cent modulation. The derived spectrum differential given in Figure 65 for approximately 100 per cent detection at a loudness level of 21 decibels would be expected to, and does, agree with the differential shown in Figure 64 for a rate of 180 rpm and a sensation level of 25 decibels. An estimated sensation level of 20 decibels within the optimal band typifies the six tests on audibility of propeller cavitation illustrated in Figures 16 through 18. The spectrum differentials observed in those tests are in good accord with expectation (see column 4 in Table 3). The components outside the optimal bands in these cases raised the total loudness to 70 phons; higher gain settings in tactical listening are uncomfortable, inefficient, and possibly injurious. Since restriction of the presentation band is unwise during search, current listening techniques provide no simple means for securing the potential improvement indicated in Figure 65. While tracking a target, bearing accuracy can probably be improved by dropping the gain or by using high-pass filters; Figure 65 implies that the latter alternative is preferable, since auditory discriminations are finer at the higher sensation levels. On the other hand, the transition curves in Figures 16, 17, and 18 show variations of slope despite the approximate constancy in sensation levels of the optimal components, and the transition curves in Figures 6 through 15 show little dependence on listening level. The factors affecting the slopes of transition curves are of practical interest, but no definite conclusions would appear warranted without further experimental investigation.

4.2.4 Supersonic Propeller Sounds

The test materials used in supersonic recognition studies were recordings of ship and submarine sounds, self-noise, and ambient noise with and without shrimp crackle. Such sound sources are commonly encountered in supersonic listening. The supersonic receivers used to obtain the reproductions were practical field installations. They admitted the frequencies between 23 and 25 kilocycles, which were heterodyned before recording, so that the recorded band, that is, the listening band, was essentially flat between 0.1 and 2 kilocycles (see Figures 66 through 75). These supersonic and audio-frequency bands are typical in supersonic listening.

Most supersonic sounds have continuous spectra, although tonal components are sometimes observed. Furthermore, the spectra of the sounds-in-the-water, where these are distributed sounds, are essentially flat over the interval from 23 to 25 kilocycles. Thus, deep-sea ambient and propeller cavitation have slopes of about −6 decibels per octave in this frequency region, but the band between 23 and 25 kilocycles extends over such a small fraction of an octave that the total change in level is negligible. The essential flatness of such spectra is not appreciably modified by selective attenuation in transmission over relatively short ranges. Similarly, the discrimination of most hydrophones against background noise changes very little between 23 and 25 kilocycles (see Section 3.3). It follows that the changes in level observed in the spectra given in Figures 66 through 75 are due principally to variations of system response with frequency.

These variations arose from two main sources. In the first place, it was necessary to use several different types of gear. Secondly, the various installations were provided with i-f and audio-frequency filters; their band widths, as well as inclusive system response over the nominal band widths, could not be very precisely duplicated from installation to installation and from time to time. Obviously, the problem of fidelity is not peculiar to the recording of supersonic sounds. It is an important problem in all recognition studies, since signal and background recorded at different times and with different installations may give misleading results if mixed together in a listening test. In connection with these test materials, it should also be noted that[7] a "limiting amplifier in the film recording process may have reduced the amplitude of the modulations which were recorded." The opinion is expressed,[7] however, that this amplitude distortion was probably not large. On the other hand, the quality of the recordings is rated no better than "fair" in most cases (see Table 5).

The apparatus used in these tests is shown diagrammatically in Figure 1. The backgrounds were heard at a level of about 60 phons, and the random method of signal presentation was followed in obtaining the results shown in Table 5, and in Figures 66 through 75. Table 5 is subdivided according to the type of masking background; effects of changing the presentation method are described in Section 4.2.6. To offset the influence of differences in the frequency response of the receivers employed to obtain the recordings, a 200-cycle high-pass and a 1,600-cycle low-pass filter were ordinarily inserted in the signal channel. This is indicated by the arrows in Figures 66 through 75, while the dashed lines show the compositions of the unfiltered signals. Use of the filters was de-

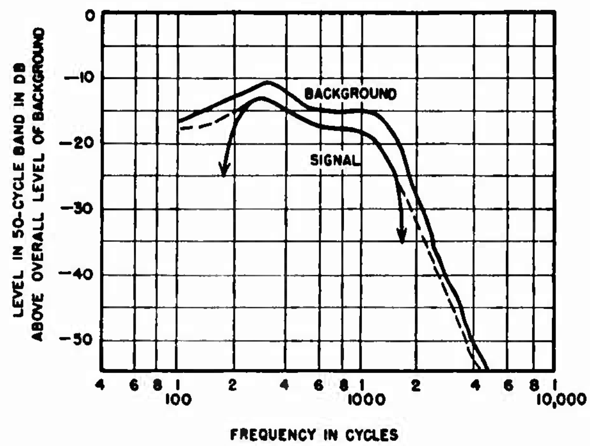


FIGURE 66. Audibility of heterodyned noise from a 12-knot transport, masked by shrimp noise; arrows show limits of the band-pass filter. *RD* for presentation band shown: —4 decibels. Character of primaudible signal: rhythm and change of quality. Primaudible frequency: 300 cycles. See Table 5, line 28. (Matched.)

signed to eliminate the "danger of the signal spectrum protruding beyond the edges of the noise spectrum, and consequently out of the masked region."[7]

The spectra give average values of rms levels in 50-cycle bands at the various frequencies; the legends in Figures 66 through 75 refer to the corresponding data listed in Table 5. To distinguish tests performed with signal and noise pairs recorded through the same and through different types of gear, the latter have been marked with an asterisk in the table and so identified in the figure captions, but, in view of the difficulty of precisely duplicating recording conditions, it is not known how closely the "matched" sounds approximate the ideal of reaching the ear over the same hydrophone.

From Table 5, the mean *RD* for the 14 "matched" sounds is −2.9 decibels, with an average deviation from the mean of 1.2 decibels. These recognition differentials apply to the entire presentation band shown in the figures. For the same group of sounds the average signal-to-noise ratio, in the optimal 50-cycle band and at primaudibility, was −1.1 decibels, with an average deviation from the

TABLE 5. Primaudibility of supersonic sounds.

Signal	Figure	Merit of signal	Character of primaudible signal	Recognition frequency region in cycles	Signal-to-noise ratio in db in a 0.1–10 kc band: 80 per cent recognition probability	Signal-to-noise ratio in db in a 0.1–10 kc band: 50 per cent recognition probability	Signal-to-noise ratio in db in the optimal 50-cycle band at primaudibility: Average ratio	Signal-to-noise ratio in db in the optimal 50-cycle band at primaudibility: Ratio for rhythm peaks	Midfrequency in cycles of optimal 50-cycle band
A. Background: recorded submarine self-noise.									
1. Destroyer Escort	....	Fair	Rough rhythm	....	−3	−5	−1	....	550
2. Twin screws of medium-size antisubmarine vessel at 13 knots	71	Good	Rhythmic chug	....	−1	−4	−3	....	600
3. Twin screws of medium-size antisubmarine vessel at 5 knots	....	Good	Rhythmic	....	−1	−3	−3	....	700
4. Aircraft carrier at 12 knots	....	Fair*	Twin screw rhythm	....	0	1	2	....	700
5. Small Coast Guard patrol boat, at 10 knots	68	Fair*	Rapid rhythm	....	−3	−3	−5	....	300
6. *E.W. Scripps* (single screw) at 7 knots	....	Fair*	Flutter	....	−2	−4	2	....	300
7. Medium size transport at 12 knots	....	Fair*	Single screw rhythm	....	−2	−5	0	....	300
8. PC at 15 knots	....	Fair*	Rhythmic	....	−2	−4	1	....	100
B. Background: recorded self-noise on patrol craft, at 10 knots.									
9. Single screw of *E. W. Scripps* at 7 knots	....	Fair	Flutter	....	1	−2	−1	....	520
10. Screws of S-class submarine at 6 knots and periscope depth	....	Fair*	Rhythm	....	−5	−7	−5	....	900
11. PC at 15 knots	73	Fair	Rhythm	....	−1	−2	−2	..	450
C. Background: recorded self-noise on patrol craft, at 15 knots.									
12. Coast Guard cutter	....	Fair*	Rhythm	....	−3	−5	1	....	1,000
13. Small auxiliary	....	Fair*	Flutter	....	−4	−6	2	....	1,500
14. Aircraft carrier at 12 knots	....	Fair*	Twin-screw rhythm	....	−1	−3	2	....	1,000
15. Small Coast Guard patrol boat at 10 knots	....	Fair*	Flutter	....	−6	−7	1	....	1,000
16. *E. W. Scripps* (single screw) at 7 knots	70	Fair*	Rapid flutter	....	−8	−10	2	....	1,000
17. Medium-size transport at 12 knots	....	Fair*	Single-screw rhythm	....	−5	−6	3	....	1,000
18. Screws of S-class submarine at 6 knots and 90-foot depth	75	Fair*	Flutter	....	−1	−3	0	....	1,500
19. *E. W. Scripps* (single screw) at 7 knots	....	Fair	Flutter	....	−1	−3	−1	....	1,500
20. Screws of S-class submarine at 6 knots and periscope depth	74	Fair*	Rhythm	....	−5	−6	−2	....	1,500

* Signal and noise were recorded through different types of gear; therefore the *RD* is unreliable.

TABLE 5. (*Continued*)

Signal	Figure	Merit of signal	Character of primaudible signal	Recognition frequency region in cycles	Signal-to-noise ratio in db in a 0.1–10 kc band: 80 per cent recognition probability	50 per cent recognition probability	Signal-to-noise ratio in db in the optimal 50-cycle band at primaudibility: Average ratio	Ratio for rhythm peaks	Midfrequency in cycles of optimal 50-cycle band
D. Background: recorded deep-sea noise. Merit: good.									
21. Aircraft carrier at 12 knots	67	Fair	Twin-screw rhythm	....	1	0	0	....	1,600
22. Small Coast Guard patrol boat at 10 knots		Fair	Flutter	....	−1	−3	0	....	300
23. *E. W. Scripps* (single screw) at 7 knots		Fair	Rapid flutter	....	−2	−4	1	....	300
24. Medium-size transport at 12 knots		Fair	Single-screw rhythm	....	−2	−4	1	....	300
25. Aircraft carrier at 12 knots		Fair	Twin-screw rhythm	....	2	0	−2	....	600
26. Small Coast Guard patrol boat at 10 knots		Fair	Flutter	....	−2	−4	−3	....	500
27. *E. W. Scripps* (single screw) at 7 knots		Fair	Rapid flutter	....	−1	−2	0	....	1,000
28. Medium-size transport at 12 knots	66	Fair	Single-screw rhythm	....	−3	−4	−2	....	300

mean of 1.1 decibels. Abnormally low recognition differentials were observed for some of the unmatched sounds, but the critical-band criterion was found to apply in such cases (see Figure 70).

These values tend to confirm the opinion of experienced listeners that supersonic signals have little character and are more difficult to detect than sonic signals; the typical cue is a rhythmic time pattern in the primaudible signal, as shown by the fourth column in Table 5. The tabulated descriptions of the primaudible signals indicate a wide variety of modulation types, but there is no corresponding variation in the primaudible signal-to-noise ratios; absence of such a correlation is discussed below. The frequencies of primaudible components were not identified with the aid of band-pass filters. From the general parallelism of the signal and noise spectra, it would be anticipated that in most cases the primaudible signals were sensed as wide-band sounds covering essentially the same frequency range as the noise. Exceptions would be expected in cases like those illustrated in Figures 70 and 75, where the spectra were not matched.

The limited information now available is inconclusive, but it seems to indicate that sonic propeller sounds are easier to detect than supersonic (see Figure 55, for example) and that the detectability of supersonic signals, in contrast to that of the sonic, is not correlated with the character of the signal modulation. It is impossible to decide, without further experimental study, whether these differences are real or accidental. A possible defect of the supersonic signals which were used, namely "clipping" of the modulation peaks during recording, has already been mentioned; also, modulation characteristics may be somewhat different at sonic and supersonic frequencies (see Figure 72 in this chapter and Figures 1 and 11 in Chapter 3). On the other hand, it is possible that supersonic listening techniques can be improved in a number of ways. Thus, the widest audio band commonly available does not exceed 2 to 3 kilocycles, because supersonic listening is currently done with receivers designed for echo ranging. Inasmuch as the echo-ranging signal is usually a tonal pulse, it is convenient to restrict the band width of the noise background so that the listening level for the primaudible frequencies may be raised to an optimal value without discomfort. On the other hand, listening to supersonic propeller sounds heterodyned from a narrow band centered at 24 kilocycles involves a number of disadvantages. High selectivity may render signal modulations less audible. Similarly, the signal envelope may have unfavorable characteristics at the lower beat frequencies; the heterodyned band may be folded, which is not necessarily an advantage, and the intensity limen is less favorable for the low audio frequencies (Figure 15 in Chapter 2). These considerations suggest that better supersonic listening might be done with a nonfolded band presented between about 0.5 and 8 kilocycles.

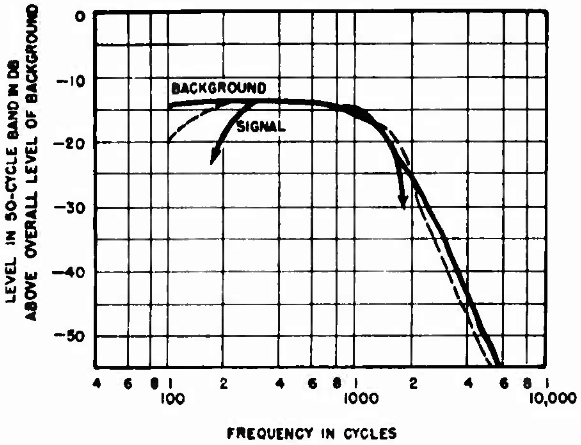


FIGURE 67. Audibility of heterodyned noise from a 12-knot aircraft carrier, masked by recorded deep-sea noise; arrows show limits of the band-pass filter. *RD* for presentation band shown: 0 decibels. Character of primaudible signal: rhythm and quality change. Primaudible frequency: 1.6 kilocycles. See Table 5, line 21. (Matched.)

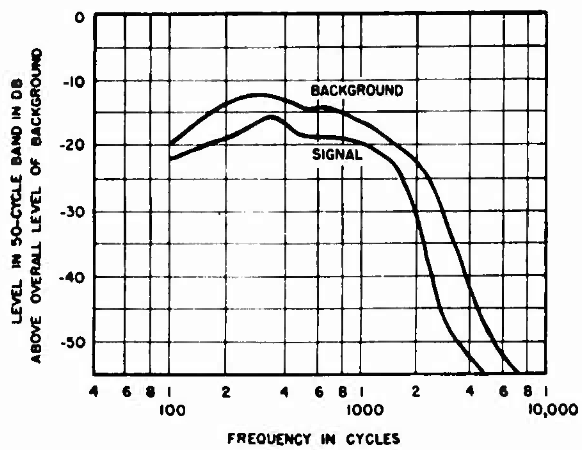


FIGURE 68. Audibility of heterodyned noise from a small 10-knot patrol craft, masked by self-noise recorded aboard a submarine proceeding at 3 knots and a depth of 90 feet. *RD* for presentation band shown: —5 decibels. Character of primaudible signal: rapid rhythm. Primaudible frequency: 300 cycles. See Table 5, line 5. (Not matched.)

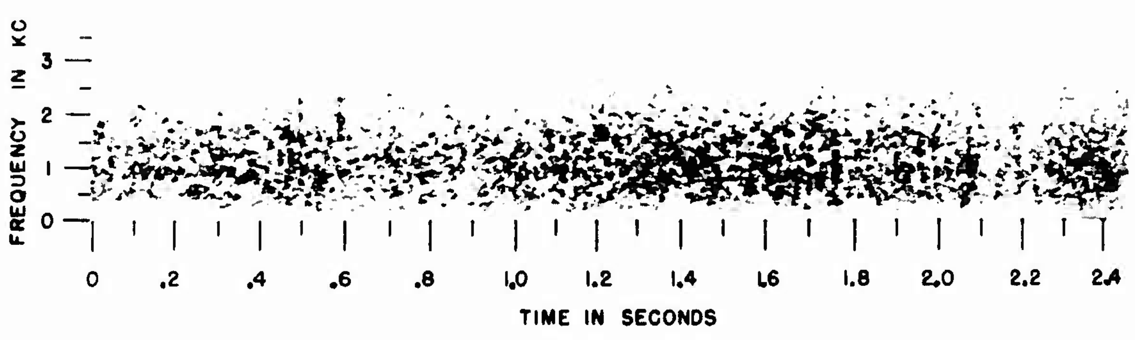


FIGURE 69. Time-frequency-intensity analysis for the self-noise shown in Figure 68.

That some such improvement may be possible is indicated by quasi-field observations[12] which show that the spectrum differential for sonic and supersonic propeller sounds alike is of the order of −6 decibels, in agreement with the components in the sonic ship signals was threshold-limited, or masking-limited (see Section 3.4.1).

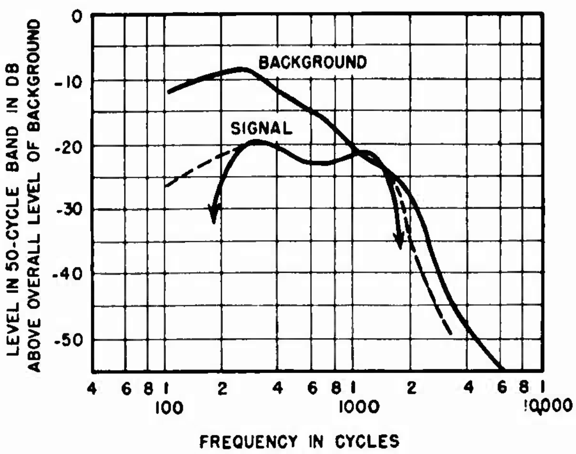


FIGURE 70. Audibility of heterodyned noise from a 7-knot schooner, masked by self-noise recorded aboard a 15-knot patrol craft at a projector bearing of 090; arrows show the limits of the band-pass filter. *RD* for presentation band shown: —10 decibels. Character of primaudible signal: rapid flutter. Primaudible frequency: 1 kilocycle. See Table 5, line 16. (Not matched.)

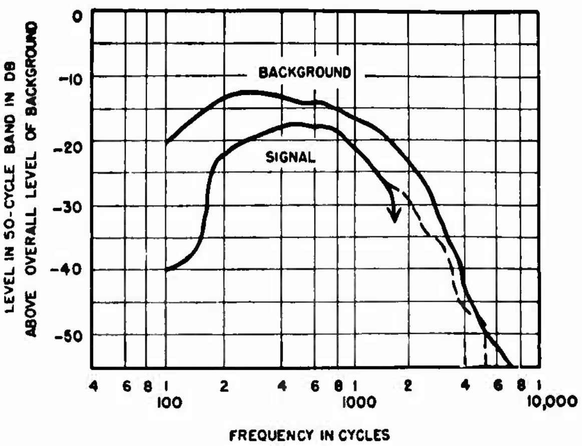


FIGURE 71. Audibility of heterodyned noise from a 13-knot Coast Guard cutter, masked by self-noise recorded aboard an S-class submarine proceeding at 3 knots and a depth of 90 feet; the arrow shows the limit of the low-pass filter. *RD* for presentation band shown: —4 decibels. Character of primaudible signal: rhythmic chug. Primaudible frequency: 600 cycles. See Table 5, line 2. (Matched.)

theory developed in Section 4.2.3. These tests were performed with simulated propeller sounds obtained from wide band thermal noise which was 50 per cent amplitude modulated.

The pure-tone threshold was determined under quiet conditions for two observers wearing headphones and is shown in Figures 76, 77, 78, and 79. When this threshold is equated to the

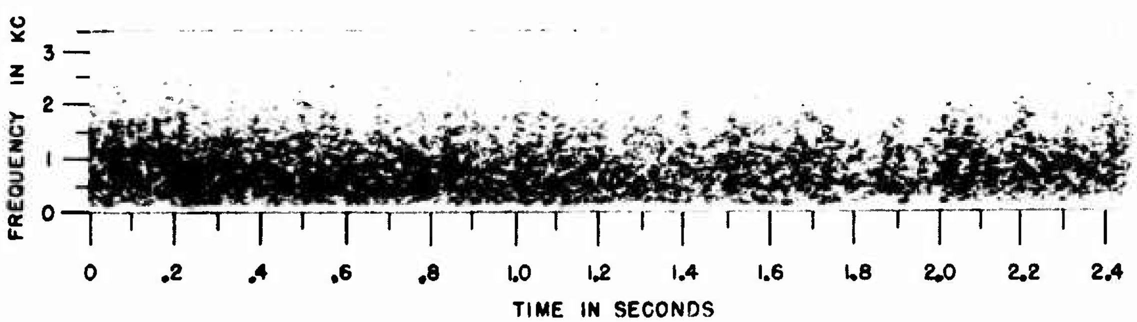


FIGURE 72. Time-frequency-intensity analysis for the self-noise from a Coast Guard cutter (Figure 71).

4.2.5 Masking of Pure Tones

Recognition tests with pure tones were conducted to study the application of the critical band criterion to backgrounds other than those used in the original work (see Chapter 2) and to determine whether the audibility of tonal free-field threshold (Figure 1 in Chapter 2) for any value above 1 kilocycle, the two sets of determinations are found to be in good agreement between 1 and 8 kilocycles, which is the highest frequency shown in Figures 76 through 79, and to diverge progressively below 1 kilocycle, so that the free-field threshold lies more than 10 decibels below the headphone threshold at 150 cycles. This inferiority of the headphone threshold at the low frequencies appears to be threshold differs from that shown in Chapter 2, since it was measured in terms of voltage applied rather than output pressure.

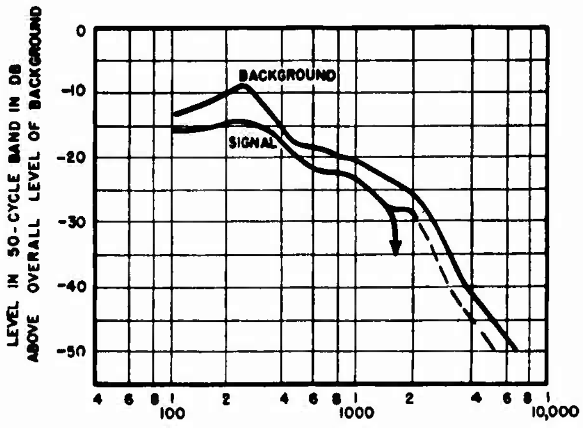


FIGURE 73. Audibility of heterodyned noise from a 15-knot patrol craft, masked by self-noise recorded aboard a 10-knot patrol craft at a projector bearing of 090; the arrow shows the limit of the low-pass filter. *RD* for presentation band shown: —2 decibels. Character of primaudible signal: rhythmic noise. Primaudible frequency: 450 cycles. See Table 5, line 11. (Matched.)

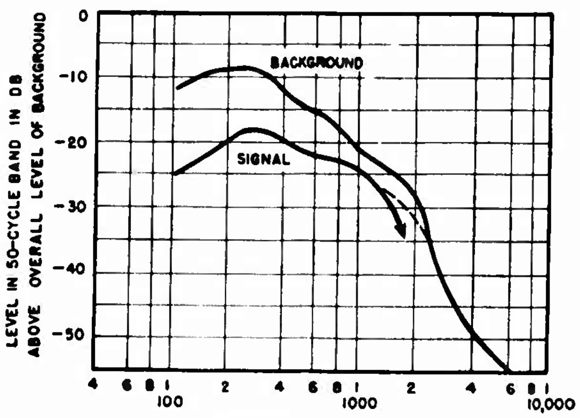


FIGURE 74. Audibility of heterodyned noise from an S-class submarine proceeding at 6 knots and periscope depth, masked by self-noise recorded aboard a 15-knot patrol craft at a projector bearing of 090; the arrow shows the limit of the low-pass filter. *RD* for presentation band shown: —6 decibels. Character of primaudible signal: changed rhythm. Primaudible frequency: 1.5 kilocycles. See Table 5, line 20. (Not matched.)

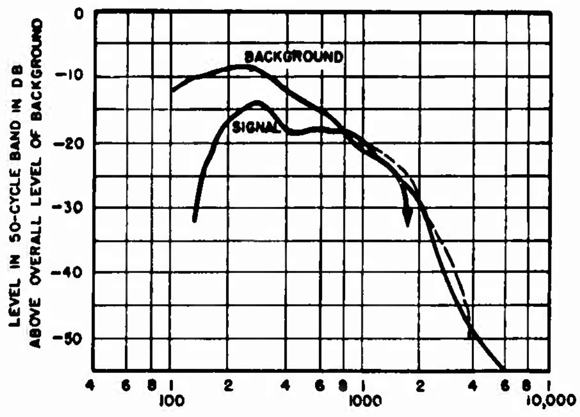


FIGURE 75. Audibility of heterodyned noise from an S-class submarine proceeding at 6 knots and a depth of 90 feet, masked by self-noise recorded aboard a 15-knot patrol craft at a projector bearing of 090; the arrow shows the limit of the low-pass filter. *RD* for presentation band shown: —3 decibels. Character of primaudible signal: roar. Primaudible frequency: 1.5 kilocycles. See Table 5, line 18. (Not matched.)

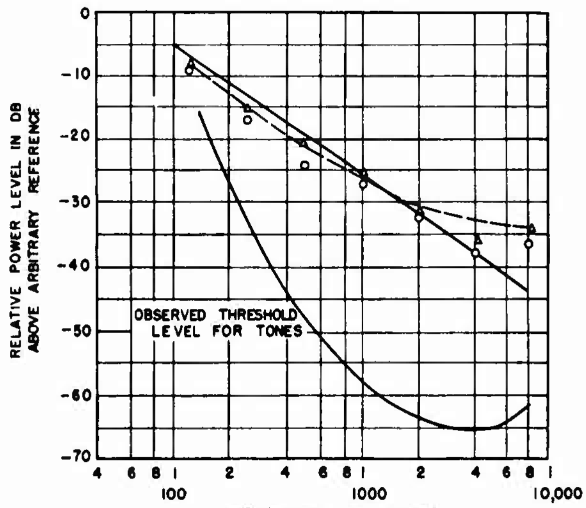


FIGURE 76. Audibility of pure tones, masked by simulated deep-sea noise. The solid line represents background measurements made with a 50-cycle filter; the dashed line represents computed levels of primaudible tones. Circles and triangles represent primaudible levels of tones determined by two observers.

typical (see Figures 76 through 79 and also Figure 4 in Chapter 10). The character at the low frequencies of the present headphone

Four widely different types of distributed noise, selected from the sounds used in the masking tests and presented over the full 0.1-

to 10-kilocycle band at a listening level of about 60 phons, were used as backgrounds. The 50-cycle spectra of these sounds are shown as solid lines in Figures 76 through 79; and critical band spectra, computed from the 50-cycle spectra by adding the quantity 10 log (critical band

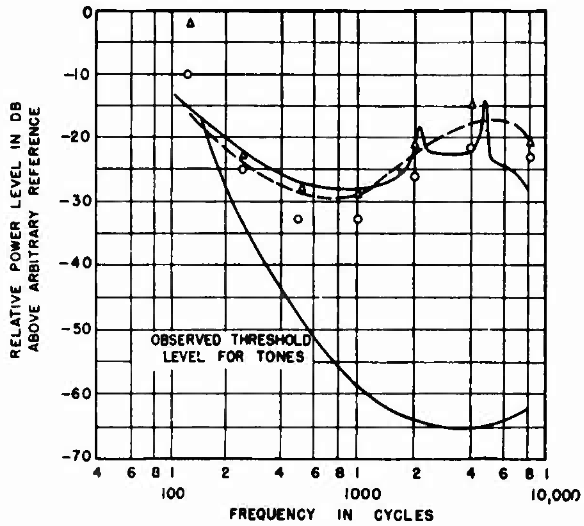


FIGURE 77. Audibility of pure tones, masked by shrimp noise. The solid line represents background measurements made with a 50-cycle filter; dashed line represents computed levels of primaudible tones. Circles and triangles represent primaudible levels of tones determined by two observers.

width/50), obtainable from Figure 17 in Chapter 2, are represented by broken lines. The two sets of spectra do not differ by more than a few decibels from 0.1 to 2.5 kilocycles.

The levels of primaudible tones at various frequencies are given as individual points, and these tone levels coincide fairly well with the critical band spectra. The performance of the two observers differs by 2 to 3 decibels, which is within the probable limits of individual variation. Error in measurement of gain setting is the simplest explanation of the apparent audibility of a 125-cycle tone below threshold (see Figure 78), and may also account for the negative tendency of the points below 1 kilocycle, which appears in this figure only. It is possible, however, that poor attenuation of room noise at the low frequencies raised the observed threshold in this region, and that airborne noise fell to a lower level during the subsequent masking tests. Changes in adjustment of the headphones may have similar effects. The 125-cycle points in Figures 77 and 79 lie 10 to 15 decibels above the critical band spectra because the level of background at this frequency is below the threshold of hearing; in other words, for optimal results, the background should be audible throughout the listening band. In general, therefore, the *RD* for a given pair of sounds will be approximately independent of gain when the signal is masked by noise received through the hydrophone. However, the *RD* will change with gain setting when signal audibility is either threshold-limited or limited by airborne noise.

The requirement that background be audible throughout the listening band may be difficult to meet if, as in the case of the shrimp background illustrated in Figure 35, the high-frequency noise components make the major contribution to the total loudness. It is possible that the deviations from the critical-band rule

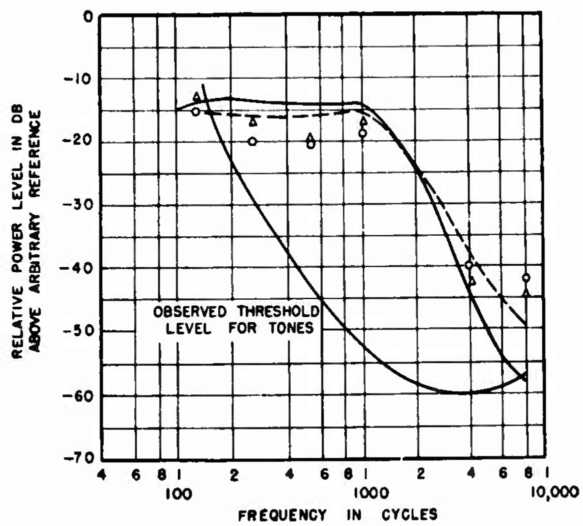


FIGURE 78. Audibility of pure tones, masked by heterodyned supersonic noise. The solid line represents background measurements made with a 50-cycle filter; the dashed line represents computed levels of primaudible tones. Circles and triangles represent primaudible levels of tones determined by two observers.

in the case of shrimp-noise backgrounds, as discussed in Section 4.2.2, were brought about by threshold-limited listening, since the background noise in those tests was also presented at a level of *about* 60 phons. However, the listening levels in the present and the previously discussed tests were only approximately equal, and in the latter case were subject to some control by the observers; hence, no definite conclusion can be drawn.

On the other hand, general use of low-pass filters in the presence of shrimp noise is inadvisable because optimal signal components at

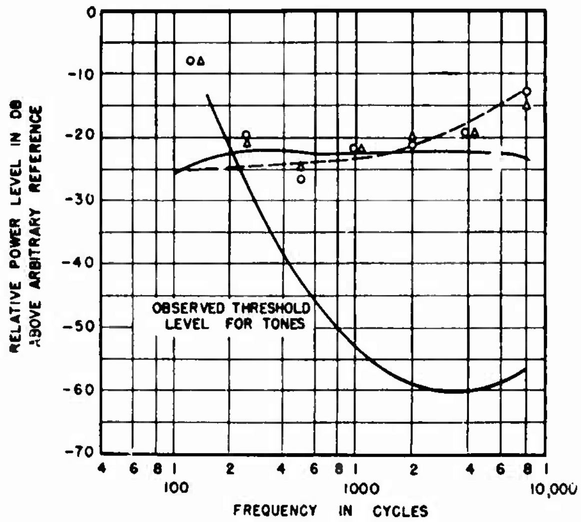


FIGURE 79. Audibility of pure tones, masked by flat thermal noise. The solid line represents background measurements made with a 50-cycle filter; the dashed line represents computed levels of primaudible tones. Circles and triangles represent primaudible levels of tones determined by two observers.

frequencies beyond the low-pass limit may be rendered inaudible. One solution would be to design the response of the entire system, including phones or speaker, so that the critical-band spectrum of the noise background, measured at the ear, parallels the threshold curve. For a hydrophone directional in one dimension (such as the JP-1), the slopes of ambient and of self-noise spectra are about 9 and 12 decibels per octave, respectively; therefore, two tilting networks would suffice to assure equally good conditions at all listening frequencies and on most occasions. Presentation like that illustrated in Figure 76 is preferable to that shown in Figures 77 and 79, since it interferes less with the audibility of low-frequency components in the signal.

The background shown in Figure 78 was obtained from a supersonic receiver, and the slope of −20 decibels per octave in the region above 2 kilocycles represents the suppression characteristic of a filter. The fact that the critical band criterion is obeyed at 4 kilocycles implies that remote masking is relatively unimportant for masking backgrounds with continuous spectra whose slope is less than −20 decibels per octave (see also Figure 6 in Chapter 2), in other words, that in such cases the masking is produced primarily by background components included within the various critical bands and that components lying outside any critical band make a negligible contribution to the masking within it. It seems probable, therefore, that the critical-band criterion also applies when the background has a slope as great as +20 decibels per octave, since remote masking is less effective in the direction of low frequencies. Finally, it may be noted that differences in time pattern of the masking backgrounds used in the pure-tone tests seem to have had little effect on the results.

4.2.6 Subjective Aspects

"In general it may be stated that learning effects were moderate in magnitude and the learning period was short. Fatigue was virtually unobserved, even for continuous listening over a 2-hour period, although this observation is not applicable to shipboard operation where the environment is much less ideal. In general the experienced listeners showed a sharper transition from 'heard' to 'not heard' than inexperienced listeners who often had a tendency to report 'heard' when no signal was present. The Sound School graduates with the highest grade on the doppler drill appeared to show a sharper transition [from 'not heard' to 'heard'] than others but they did not appear able to hear a [masked] signal at a lower [relative] level than others."[7]

The doppler drill, which is a standard test used in selection of sonar personnel, measures ability to discriminate the pitch of a test tone from that of a reference tone. The tones are presented in succession and without interference from masking sounds. Absence of correlation between performance in the doppler drill and the masking tests is surprising in view of the close relation between frequency discrimination and critical band width (see Sections 2.2.1 and 2.3). This point may deserve further study; for example, the observer who gave consistently better-than-average performance in the pure-tone tests (Figures 76 through 79) was a trained musician.

A set of tests conducted in connection with the proposed masking monitor provides a useful index to the typical spread among recognition differentials for individual listeners. Masking tests were made in the laboratory using a recording of a fleet-type submarine at periscope depth and a speed of 3 knots. Nine listeners participated; of these, only three had previous experience in masking tests. In this test, the listeners could control the level of the masking background, which was wide-band thermal noise, and the gain setting in the signal channel was fixed at a comfortable listening level. "All [observers] were given the same instructions," according to an informal communication from the University of California Division of War Research [UCDWR], "to turn the noise dial until the submarine sound was just masked by the noise, then to back it off until the sound was again audible. The noise attenuator values at both conditions were recorded and the mean taken as the recognition value.

"In the first test, the determinations were made without interrupting the submarine sound. In the second test, the subjects were instructed to interrupt the submarine sound at will, to [ascertain] whether such technique facilitated determinations." The results are shown in Table 6.

TABLE 6. Effect of experience and signal modulation.

	Signal not interrupted Masking-noise level in db			Signal interrupted Masking-noise level in db		
	Signal audible; noise decreasing	Mean value	Signal masked; noise increasing	Signal audible; noise decreasing	Mean value	Signal masked; noise increasing
3 different experienced observers	12	12.5	13	15	15.5	16
	14	15	16	16	16.5	17
	14	15	16	16	18	20
Average in db	....	14.2	....	....	16.7	....
Mean deviation in db	....	1.1	....	....	0.9	....
6 different inexperienced observers	6	8.5	11	16	17	18
	9	9.5	10	14	14.5	15
	12	13.5	15	18	18.5	19
	13	14	15	15	16	17
	10	11	12	15	16	17
	12	14	16	14	15.5	17
Average in db	....	11.8	....	....	16.2	....
Mean deviation in db	....	2.1	....	....	1.0	....
Difference in db between best and poorest observer	....	6.5	....	....	4.0	....

All levels in Table 6 are stated in decibels relative to a single arbitrary reference, and the higher noise levels correspond to better performance; thus the table shows relative, not absolute performance. The method of measuring performance in this case was essentially that of changing the nature of the stimulus in a progressive series of small steps which traversed, and thereby defined, the region of the response threshold. It is known as the *method of minimal increments* and has the advantage of being less time consuming than the random-order method. The method of increments defines the borders of the transition zone but gives no clear picture of the transition curve. It is commonly observed in such tests that the observers' thresholds are biased, in the sense that performance is better when the stimulus is changed from "heard" to "not heard" (labeled "noise increasing" in Table 6) than when the reverse change is made ("noise decreasing"). This is often attributed to a "memory effect"; in other words, if the cue is definite at the outset and is gradually faded out while the observer continues to listen, he can usually follow it a little further below noise than when the cue is not sensed in the beginning and subsequently emerges into consciousness. This type of response probably depends more on continuity of sensation than it does on knowing and remembering the character of the primaudible cue. In practice, therefore, it is generally easier to maintain contact with a target than to make first contact.

The mean of the ascending and descending threshold determinations is usually taken as the minimal-increment *RD* (for an individual or a group); and the mean *RD* in the UCDWR tests was usually about 2 decibels larger (less favorable) than when the determination was made by the random-order method. Stated differently, a minimal-increment determination corresponds approximately to the 75 per cent point on a recognition probability curve, since an increase of about 2 decibels will, on the average, raise the recognition probability from 50 to 75 per cent (see equation (4) in Section 4.1.4).

It has been noted in Section 4.1.5 that the opposite result was apparently obtained by the British, who found that the minimum-increment recognition differentials were frequently less (more favorable) than the random-order recognition differentials, some differences amounting to as much as 5 to 6 decibels. This finding may be attributed to differences in indoctrination used in the two sets of tests. In the British random-order tests, the observers were not trained to listen consciously to sounds of all frequencies; in the random-order tests they presumably concentrated on that region of the spectrum where the target sound was most audible in the absence of the masking background, while in the increasing-level tests they were forced by necessity to listen to all frequencies, since the target sound had not yet been heard. The UCDWR observers were trained in all cases to scan the frequency spectrum and were warned that the most prominent features of the signal at a high sensation level might not be the cues at primaudibility.

Table 6 also reveals the effectiveness of auditory motion, whether introduced by modulation of the signal or by training a directional hydrophone. Thus, interrupting the signal improved the average performance of the experienced group by 2.5 decibels and that of the inexperienced group by 4.4 decibels. In other words, when rapid comparisons can be made between pure background and the signal-background mixture, the minimal-increment *RD* corresponds roughly to 50 per cent, rather than to 75 per cent recognition probability. It will be noted, incidentally, that signal interruption affects both ascending and descending recognition differentials, that is, the entire transition curve is shifted. The communication describing these results observes that "although some of this improvement might have resulted from experience gained in (the first test), the comments of the observers indicated a greater sense of reliability for the signal-interruption type of test."

The relative performance of the two groups was nearly identical under the second condition of test, but even in this case, the best observer outperformed the poorest by 4 decibels. The practical situation is more complicated than the one used in this test; hence, the effects of training and aptitude will usually be larger. Table 6 is in agreement with previous indications that signal modulation may help by 2 to 3 decibels (see Section 4.2.2); possibly this factor obscured such differences in performance as would be expected purely on the basis of grades in the doppler drill.

For experienced observers, the average deviation from the mean performance in the minimal increment tests was about 1 decibel. This was also true of the various random-order tests, except for cases in which different components in the signal were detected by different listeners. Under these circumstances, discrepancies as great as 8 decibels occurred (see also Section 4.1.5). In cases like that illustrated in Figure 34, some observers reported the signal primaudible at 120 cycles; some, at 1,400 cycles; and others, at both frequencies.

In general, it was found that judgments were easiest and most consistent when a hydrophone sweep was simulated by fading the signal into background and out again. The random-order method was of intermediate difficulty, since it requires that a mental image of the pure background noise be retained. The judgments were most difficult to make and the response least consistent when the observers listened to the background noise continuously for a 2-hour period during which various signals were *slowly* brought into audibility at random intervals (separated by about 10 minutes, on the average) and faded out again. Detection under these conditions required, on the average, about 2 to 3 decibels more gain in the signal channel than was needed for 50 per cent detection of the same signals by the random-order method, which corresponds roughly to the 90 per cent point on the transition curves. Auditory fatigue was apparently not an important factor in these tests, since the observers performed as well at the end of the two hours as at the beginning. The 2-hour test resembled shipboard listening in some respects, but it gave no weight to the help which would probably be derived from training a directional hydrophone. For practical purposes, therefore, the 50 per cent *RD* determined by the random-order method is probably as useful an index of detectability as any.

4.3 CUDWR-NLL TESTS

In addition to the field tests described in Chapter 5, some laboratory tests on masking of sounds were also carried out at New London. These tests were roughly similar in technique to those described in previous sections of this chapter. Their objectives were somewhat less general, however, in that they were designed to give specific answers to certain questions which arose in the design of sonic listening gear. These tests are described in the next three sections.

4.3.1 Interval Tests

A study[14] was undertaken which was designed to determine whether "an optimum length of listening interval exists when switching from one hydrophone to another in a series of cable-connected hydrophones or from one buoy to another in the case of radiosonic buoys." To this end, masking tests were made in which various signal-background mixtures were presented to the observers for selected lengths of time. As shown in Figure 81, performance was significantly influenced by the length of the listening interval.

Four test series were run, most of them in duplicate. In each series, 25 pairs of sounds were presented according to the time schedule shown in Figure 80, where A and C represent

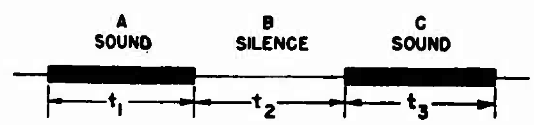


FIGURE 80. Sequence within a test pair, where time intervals t_1, t_2, and t_3 are equal.

members of such a test pair, each member being equal in duration to the intervening period of silence denoted by B. When the first of the three intervals shown in Figure 80 consisted of a mixture of signal and background, the third contained only background, and vice versa. The order in which the mixture and the pure background occurred in successive pairs, as well as the signal-to-background ratio for the mixture, was randomized, but the lengths of the equal time intervals (t_1, t_2, and t_3) were restricted in each sequence of 25 pairs of sounds to 2, 3, 4.5, or 7 seconds. Between successive pairs in a test sequence of 25, the observers indicated on a score sheet whether, in their judgments, the signal had occurred in interval A or interval C of the immediately preceding test pair.

Since the experimental variable of interest was the length of listening interval, the same signal and background were used throughout. The former was a recording of the underwater sonic output of a freighter and contained a typical assortment of thumps, engine noises, and propeller sounds; the masking background was water noise, extending over a wide frequency band. The apparatus used was essentially that illustrated in Figure 1, except that a wide-band high-fidelity loudspeaker was substituted for the headphones. The gain in the background channel was set and maintained throughout the test at a fixed level which the listeners found comfortable. The signal-to-background ratio for successive pairs was varied at random over the range shown in Figure 81, which was covered in 5-decibel steps as shown by the experimental points. Since there were 11 observers and duplicate tests were made for all but one of the signal-to-noise ratios, the experimental points corresponding to −5 decibels represent 55 independent judgments; all other points in Figure 81 represent 110 judgments. The signal-to-noise ratio labeled 0 decibels in this figure represents the value which experienced observers designated as primaudible. This determination of primaudibility was made prior to the interval tests by the method of minimal increments, and no time limitation was imposed on the listeners. The various signal-to-noise ratios specified in Figure 81 are referred to this independently determined minimum increment *RD*.

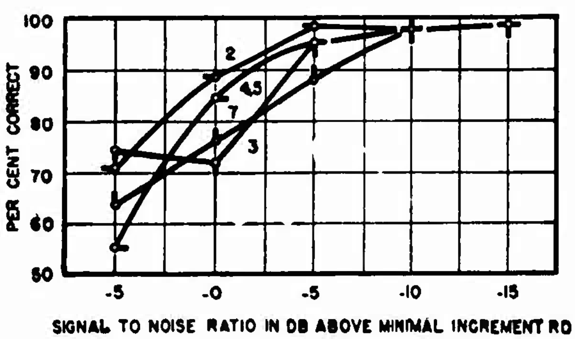


FIGURE 81. Effect of presentation interval on detection, where numerals indicate durations of presentation intervals in seconds.

In essence, the four sets of points which have been connected by curves in Figure 81 are points defining transition curves. These transition curves differ in a number of ways from the ones given in Figures 6 through 23. For example, in the present case the ordinate indicates primarily the relative number of successful assignments of the signal to the proper member of a test pair; since the subjects are required in this type of test to make some assignment for each pair of sounds, they should in general receive a score of 50 per cent for signal-to-noise ratios at which the probability of *perception* is vanishingly small. In other words, a score of 50 per cent corresponds to inability to make the required discrimination, or 0 per cent perception. Similarly, a score of 75 per cent in this type of test is usually considered equivalent to a 50 per cent probability of perception. Setting up these correspondences between p per cent on the curve of perceptions and c per cent on the curve of correct answers may be expressed quantitatively in the following way. The quantity $(100-p)$ per cent represents the "recognition impairment," the relative number of presentations of a given kind in which the observer fails to make the desired discrimination. If he guesses in these cases, he will in the long run make the correct guess half the time. Hence, the apparent impairment $(100-c)$ per cent will be only half as great as the true impairment $(100-p)$ per cent. Thus, $(100-c)/(100-p)=\frac{1}{2}$, and $c=(100+p)/2$. Substituting in this expression shows that, as already stated, the following sets of values of p and c are equivalent, and similarly at intermediate points:

p per cent	100	80	50	20	0
c per cent	100	90	75	60	50

From equation (4) in Section 4.1.4 it will be seen that the half-spread, or decibel increase in the signal-to-noise ratio which improves the probability of perception from 50 per cent to 80 per cent is $6/n$ decibels. The preceding tabulation shows that an equivalent number of decibels separates the 75 and 90 per cent points on the curve of correct responses. The interval between 75 and 90 per cent averages 4 decibels for the curves in Figure 81 whence, $n=1.5$ decibels. Since a decrease of $6/n$ decibels produces a 30 per cent impairment in the probability of perception, the ratio $(6/n)/30$ states the number of decibels required to produce a 1 per cent impairment (over the nearly linear portion of the transition curve between the 20 and 80 per cent points). When $n=1.5$, the decibel loss which leads to a 1 per cent impairment is nearly twice as large as in the typical case (where n is about 2.5). This is due, presumably, to the fact that observers usually fail to report the doubtful perceptions when guessing is penalized, as in the random-order tests discussed earlier in this chapter. In practice, therefore, where doubtful contacts are reported as a matter of course, the spreads of the relevant transition curves are probably somewhat larger than shown in Figures 6 through 23, and more nearly those in Figure 81. It should also be noted that the curves in Figure 81 do not approach a value of 50 per cent even for the smallest signal-to-noise ratio shown; that is, some discrimination is possible at ratios more than 5 decibels below the nominal threshold value, which in this case occurs at the 75 per cent point (see also Figures 6 through 23). Similarly, it will be clear from Figure 81 why an interval of 2 decibels is usually considered necessary for precise definition of the transition curve.

It proves helpful to consider an additional difference between the procedure used in the tests described in Sections 4.1 and 4.2 and the procedure used in these interval tests. In the present case. the observer is able to make a more nearly immediate comparison between the mixture and the pure background than is possible in the former, and is, therefore, less dependent on a memory image of the primaudible cue. Such images fade rather quickly. Hence, the fact that two members of a test pair could be compared after a smaller time lapse for the shorter listening intervals (see Figure 80) may well have had more weight in determining the nature of the results than did the durations of the test sounds themselves. Thus, it will be seen from Figure 81 that in general — and aside from irregularities like the one in the curve relating to the 3-second interval (which irregularities could probably be minimized by making a larger number of determinations and varying the signal-to-noise ratio in smaller steps)—the trend associated with a decrease in the time lapse between members of a test pair is an increased ability to perceive the signal. This improvement would be expected on the basis of the discussion in Sections 5.1 and 5.6. It may be noted that the 75 per cent point on the curve for a 7-second interval (an interval between successive presentations which is nearly equal to that employed in the previously described test procedures) falls on the 0-decibel line, which represents the minimal increment *RD* for experienced observers, whereas the 75 per cent points for shorter time intervals occur at smaller signal-to-noise ratios.

In assessing the difference of nearly 5 decibels between the 75 per cent points of the curves for 2 and 7 seconds, it should be noted that impairment of perception due to unfavorable coincidences between signal and noise peaks is probably not very significant for listening intervals more than 1 second in duration (see Section 8.4.4). Operation of this factor furnishes an additional reason for sweeping the axis of a directional hydrophone several times in succession across a suspected bearing in order to obtain the maximum advantage from the locally produced modulation. On the other hand, the rapid pace and higher degree of concentration required when the listening interval was reduced to 2 seconds was considered too fatiguing by the observers, who agreed that periods of active listening should probably be no less than about 3 to 4 seconds during a watch of practical length. Hence, the presentation interval of 5 to 10 seconds normally used in listening tests is probably adequate so far as reduction of fatigue effects is concerned. Some observations made during British field tests of radiosonic equipment indicate that it is probably not useful to extend a watch beyond a half hour; at the end of that time, listeners tended to become confused and to make increasingly numerous errors of commission. Finally, it should be pointed out that the 7-second interval was the one used in the first test of this series. Hence, if any learning occurred during the course of the tests it probably affected the results more for the 7-second interval than for the others; such an effect would make the observed improvement above performance in the 7-second test greater than the improvement actually due to diminished time lapse.

4.3.2 Peak Tests

The response of hydrophones often shows one or more peaks, or maxima, that occur at frequencies to which the hydrophone structure resonates. When such resonances fall within the intended listening band, they may limit the usefulness of the receiver. If the resonance frequency, or frequencies, cannot conveniently be shifted beyond the confines of the listening band, it may still be possible to control their magnitudes by relatively simple means. Thus, it becomes desirable to know how high a response peak may be tolerated without significantly impairing performance. The tests described in this section[15] were designed to determine the permissible heights of resonance maxima in the response of hydrophones like the JP-1. The latter, a 3-foot nickel tube, may show a disturbing degree of longitudinal resonance (vibrations propagated along the axis of the cylinder) unless its ends are clamped between rubber blocks so as to damp out this mode of vibration.

It may be noted in passing that transfer of energy to the longitudinal mode of vibration is unavoidable since the application of stress along a given axis of a solid produces a deformation not only in that direction but also along axes transverse to the first; the magnitude of the secondary strain in terms of the primary is given by Poisson's ratio. If the longitudinal vibration of a tubular hydrophone supported along a transverse axis through its center resembles that in a rod clamped at its center, the fundamental mode of vibration involves a wave whose length λ is twice that of the cylinder axis d; in other words, $\lambda = 2d$, or $f = c/2d$, where f is the resonance frequency and c is the velocity of sound propagation (about 1.6×10^4 feet per second in the case of nickel). For a 3-foot nickel tube, the fundamental frequency of the longitudinal resonance should therefore occur at about 2.5 kilocycles, which is within the sonic listening band. Although such a peak is not shown in Figure 2 of Chapter 5, because the resonance frequency was not used during calibration, fairly strong maxima are often observed in the spectra of self-noise received in JP-1 gear.

In practice, the hydrophone is immersed in a random noise field (this description applies to most signals as well as most backgrounds); hence, the resonance vibrations are shock-excited intermittently and persist, or decay, over an interval which depends on the "Q," or sharpness of tuning, of the resonance characteristic. Thus, when a broad-band sound is received over a resonant hydrophone, the listener hears a fluctuating tonal ring (like that emitted by a public address system with a moderate amount of acoustic feedback) superposed on the sound proper. When the ringing sound is a prominent and fairly continuous component of the normally received background, the resulting annoyance and distraction interfere with careful search and may lead the operator to drop the gain too far. If the optimal component received from a given target lies in a frequency region fairly close to that of the resonance peak, the signal may not become primaudible until the signal-to-noise ratio is significantly greater than would otherwise be required. This would not necessarily be true, however, especially if the time pattern of the resonance is different for the mixture and for pure background. For example, the resonance peak shown near 2 kilocycles in Figure 77 apparently had very little effect on the audibility of the 2 kilocycle tone.

The present test sheds no light on the masking problem, but it does indicate the condition in which a resonance peak will be annoying. Since such peaks will not annoy unless they can be heard, the observers were tested on their ability to detect the presence of resonance peaks in several sequences of sounds. Two recordings were used: one obtained from the sonic underwater output of a small surface vessel, and the other, from that of a surfaced submarine. The recording hydrophone, system, and medium were fairly flat in frequency response and relatively free of resonances. Thus, the outputs from the recordings were wide-band sounds whose spectra had the characteristic negative slope of about -6 decibels per octave. In order to simulate the rising response characteristic of the JP hydrophone (see Figure 2 in Chapter 5), the electrical outputs from playbacks of the recordings were transmitted through a tilting network which increased the applied gain at the rate of about 6 decibels per octave (see Figure 82), so that the presented sounds had essentially flat spectra. This transmission network also contained controllable resonance elements by means of which it was possible to simulate the intermittent ringing note due to hydrophone resonance and to bring the resonance peak in at any of three different frequencies (0.9, 1.4, and 2.0 kilocycles), with one of six heights relative to the adjacent spectrum levels (the added gains were 0, 2, 4, 6, 9, and 12 decibels), and with a Q factor of 30, 35, or 50. One of these conditions is illustrated in Figure 82 where the dotted portion of the curve applies to a peak height of 0 decibels.

The test sounds were presented to between 9 and 11 observers by means of a high-fidelity loudspeaker and at a comfortable listening level. No sound mixtures were used, and the only distinction that the observers had to make was between presentations in which the resonance note was audible and those in which it was inaudible. Ten random-order tests were administered, each consisting of 24 sound samples among which the six available peak heights were equally and randomly distributed; samples were 5 seconds long and were separated by a silent interval of 3 seconds during which the listeners expressed their judgments on prepared score sheets. The ship sound recording, as well as the frequency and sharpness of the resonance peak, was fixed throughout each group of 24 samples, the only variable being peak height. Among the ten tests, however, different combinations of recording, Q, and peak frequency were tried. For each of these ten combinations, typical random-order transition curves were obtained, showing the probability of detecting the presence of the resonance note as a function of peak height.

In most cases, a peak height of 6 decibels was recognized in 50 per cent of the trials, and a height of 9 decibels was detected in 80 per cent of the trials (in other words, $n=2$). Two experienced sonar operators, who did not participate in the preceding tests, indicated that they found a peak of about 8 decibels annoying. The transition curves showed almost no dependence on the ship sound used or the frequency of the peak (for the three frequencies tried), but were significantly influenced by the Q of the resonance so that lower peaks were more readily detected when Q was large. Among other things, the duration of the intermittently excited ringing note is determined by the sharpness of tuning (see Section 7.1); for the tests described here, the durations of the tonal pulses which constituted the audible evidence of a resonant condition were of the order of 0.1 second (as indicated by the widths of the resonance peaks). In Section 2.2.2, it was shown that the subjective loudness of a tonal pulse, as well as its audibility in the presence of masking background, is less than that of a sustained tone of the same frequency. Figure 8 in Chapter 8 shows that the loss in audibility of a 100-millisecond pulse, relative to that of a sustained tone, is of the order of 7 decibels, and that this loss is comparatively independent of frequency over a large frequency interval. In other words, a mixture of pulse and noise background can be distinguished half the time from the pure background (contained in a critical band centered at the pulse frequency) when the pulse-to-noise ratio is about 7 decibels and the pulse length is about 100 milliseconds; essentially these conditions prevailed during the present tests on the audibility of resonance peaks. Thus, it may be inferred from the preceding observations that a resonance note will be annoying whenever it can be heard, but that its audibility is diminished when its duration is brief. Therefore, damping a system resonance will improve hydrophone quality by increasing the broadness of the peak as well as by limiting its height.

4.3.3 Frequency Translation of Airplane Noise

It would be tactically valuable to submariners preparing to surface if they could detect the presence of hostile aircraft by means of the radiated noise and at ranges corresponding to at least 30 seconds of flying time. Since this type of signal is characterized by a number of prominent single-frequency components at the very low sonic frequencies (where critical band width is less favorable to detection than at higher frequencies and where, moreover, the sensitivity of the ear and the response of headphones are poor), a study was made[16] of the possibility that performance could be improved by heterodyning the received signal and thereby shifting it to a more favorable frequency region.

Recordings were made of the noise received from an airplane in level flight, and following a straight course, as it approached and then receded from a microphone supported just above the sea surface. Since wind noise, due to the flow of air about the receiver structure, and the splashing of waves masked the airplane signal at each end of the run, the length of time during which the signal-to-noise ratio was sufficient to assure signal audibility provided a convenient measure of comparative performance. The mean intervals of audibility were determined for the same group of observers listening to a playback of the sounds as originally recorded, and then to the output obtained by frequency-translating these sounds either 200 or 500 cycles. In some of the tests, both side bands were presented to the observers; in others, one of the bands was eliminated by means of filters. (When any signal component is heterodyned, the resultant signal consists of two side bands; one representing the sum of, and the other the difference between, the frequencies of the original component and the heterodyne note.)

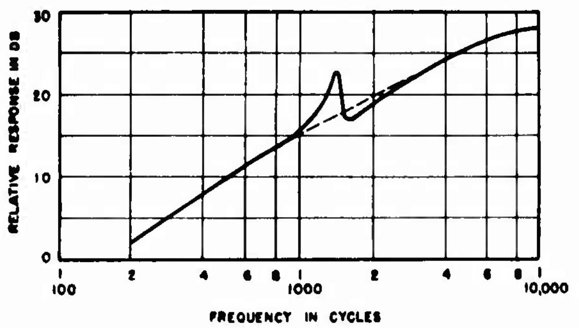


FIGURE 82. Frequency response characteristic of system used in peak tests. Peak had a frequency of 1.4 kilocycles and a height of 6 decibels, with a Q of 35.

No significant quantitative difference was found to exist between the intervals of signal audibility in the heterodyned and unheterodyned recordings. However, the listeners found the heterodyned sounds qualitatively dissatisfying, apparently because the uniform frequency shift imposed on each of the signal components destroyed the nearly harmonic relation between them which occurs in the sound as originally received, and consequently destroyed much of the character typifying the original signal, thereby increasing its resemblance to the random noise background. In other words, there seemed to be no actual decrease in ability to discern the presence of the wanted sound, but less subjective certainty that the sensed cue was really the sound wanted. Precautions had to be taken during tests with the heterodyned sounds to eliminate circuit hum, as well as rumbles generated by motion of the turntable; that is, these sounds were below the audibility threshold when the observers listened to the unheterodyned recording.

It would appear to follow that the optimal component was not threshold-limited for the sounds as originally recorded; on the other hand, it is uncertain whether the fidelity of this recording, in the very low-frequency region, was adequate to the purpose at hand. One further possibility is worth considering: that a component with a very favorable signal-to-noise ratio, inaudible in the unheterodyned recording because of threshold limitation, was rendered audible by frequency translation, but that the improvement which might be expected from this effect was offset by the accompanying destruction of harmonic relations and the help they sometimes afford in signal detection (see Section 8.5). In further studies along these lines, it may be profitable to multiply the component frequencies by a constant factor, thereby preserving the harmonic relations instead of translating them all by the same amount.

It may be mentioned, in conclusion, that the underwater radiation from ships is generally rich in prominent single-frequency components in the subsonic range (1 to 10 vibrations per second) which are associated with various modes of hull vibration. The question whether this region could profitably be exploited with listening gear designed for long-range detection has apparently received no systematic study.

Chapter 5

FIELD MEASUREMENTS ON MASKING OF TARGET SOUNDS

A GROUP OF PRACTICAL surveys, made under field or semifield conditions, have been described by the New London Laboratory of Columbia University Division of War Research (CUDWR-NLL).[1-11] These surveys were undertaken in connection with the design, development, and evaluation of practical hydrophone installations, and they furnish useful information which is not obtainable from tests of the kind already discussed. However, a certain amount of caution is required when examining the results of field studies since precise acoustic measurements are more difficult to make under sea conditions than in the laboratory. In addition, it is difficult to evaluate, control, and reproduce the experimental factors, to vary them one at a time, and to assure that the range of the variables, known and unknown, which has been sampled is either representative or adequate. In consequence, the quantitative significance of numerical data is often open to question and interpretation.

The objective of these tests was to study the controllable factors which affect the detection ranges and bearing accuracies of sonic and supersonic installations. However, the methods used and the results obtained in these supplementary studies have considerable intrinsic interest, and their discussion permits the fundamental problems to be formulated in more concrete terms. The observations on primaudibility, that is, on detection range, are summarized in the present chapter. The experimental variables were hydrophone type, listening band, and auditory motion produced by training the listening hydrophone or changing the signal level.

5.1 HYDROPHONE DIRECTIVITY IN SONIC MASKING

Before examining the data, it will be useful to summarize and extend the earlier discussion of hydrophone directivity given in Section 3.3 and to correlate some of the properties of such devices with those of the ear. Hydrophones discriminate against nonaxial sounds; in other words, responsiveness depends upon angle of incidence, and the width of the main lobe (that angular interval in which response is relatively high) is a measure of the discrimination against nondirectional sounds. This lobe width, in radians, is equal to $2\lambda/d$, where λ is the wavelength of the sound and d is the linear dimension of the hydrophone, provided λ is much less than $d/2$.

For incident sounds whose half wavelength exceeds the dimensions of the hydrophone, directivity becomes negligible. Hence, no effective discrimination against a nondirectional sound field, such as sea noise, can be obtained when the half wavelength exceeds the major linear dimension of the hydrophone, since there will be no significant difference of phase across the hydrophone. Thus the condition for directivity is $\lambda > 2d$. On introducing the relation $\lambda = c/f$, where c is the velocity of sound, this condition becomes $f > c/2d$. The velocity of sound in sea water depends on temperature, salinity, and hydrostatic pressure; a useful average value is 4,800 feet per second.

To be more specific, consider the JP-1 hydrophone, a type frequently used in the New London tests. The code designation JP-1 refers to a hydrophone fashioned from a hollow nickel cylinder about 3 feet long and 2 inches in diameter. The magnetostrictive nickel tube completely encloses a copper coil whose terminals are led, through the wall of the tube, to a listening amplifier. At the coil terminals appear electrical variations induced by the changes in magnetic flux set up by incident sound. The small diameter of the JP-1 hydrophone precludes directivity for underwater sounds incident in the plane perpendicular to the geometrical axis of the cylinder and with frequencies below 14.4 kilocycles ($c/2d = 4{,}800/0.33 = 14.4$ kilocycles).

In the plane containing the geometrical axis of the cylinder, there is useful directivity between 5 and 10 kilocycles. But even in this plane, the discrimination of the bare hydrophone is negligibly small (less than 2 decibels) for frequencies below 1.3 kilocycles ($c/2d = 4{,}800/6 = 1.3$). With a baffle, which shields half of the hydrophone surface, the discrimination against nondirectional sound is approximately 3 decibels greater, at all frequencies, than is shown by Figure 18 in Chapter 3 (see also Figure 2 in this chapter). From the preceding discussion, it is clear that the JP-1 hydrophone does not have a unique axis. Its three-dimensional directivity pattern, when the bare hydrophone is immersed in a nondirectional sound field, has been described as resembling "a washer with the hydrophone stuck through it like a bolt."

The JP-1 hydrophone was standard sonic listening gear on submarines and inshore patrol craft. The JP-1 is generally mounted on the foredeck of submarines, where the conning tower serves to shield it from the vessel's own propeller sounds; in addition, the submarine may bottom without risking injury to the hydrophone. The geometrical axis of the JP-1 is rotated about its midpoint and in the horizontal plane. In other words, the plane of the "washer" can be trained to different compass points; and the absence of directivity in the vertical plane is helpful in maintaining contact with surface targets at short range. When mounted on surface vessels, the JP-1 is generally suspended through a well in the hull and is then described as *through-the-hull* [TTH] gear. The magnetic shielding of this hydrophone is incomplete when it is mounted on surface vessels. Owing to the high level of ignition noise from propulsion machinery and excessive stresses in the supporting shaft due to hydraulic drag at the higher speeds, TTH gear is usually drawn up against the hu'l and secured with its length parallel to the keel when the craft is underway.

It will be noted from Figure 18 in Chapter 3 that the discrimination of a line hydrophone against isotropic sound improves at a rate of 3 decibels per octave for frequencies greater than $c/2d$. This behavior follows from the fact that the width of the main lobe is inversely proportional to fd; when d is fixed and f is doubled, $10 \log (2fd/fd) = 3$ db. (The ordinates are negative in Figure 18 of Chapter 3 to show the decrease in the level of received sound at higher frequencies.)

The sensitive surface of the type of hydrophone commonly used for supersonic listening and echo ranging may be approximated, for purposes of this discussion, by a circular disk 15 inches in diameter (see Figure 17 in Chapter 3, and Section 3.3). The receiving disk is suspended along and rotated about a vertical diameter; it is usually protected from the sea by enclosure within a sealed liquid-filled metallic shell or streamlined dome which is mounted at keel depth outside the hulls of surface vessels and bottomside on submarines.

While the horizontal and vertical lobe widths of a line hydrophone are unequal, leading to the "washer" pattern, in the case of a circular disk the hori tal and vertical widths are equal. The spatial configuration of the main lobe for a disk hydrophone is that of a cone with its axis perpendicular and its vertex adjacent to the hydrophone face. In other words, the disk differs from the line in that it discriminates against sea noise in the vertical plane as well as the horizontal. The discrimination should therefore be inversely proportional to $(fd)^2$ and should improve at a rate of 6 decibels per octave at frequencies in excess of 1.9 kilocycles ($c/2d = 4{,}800/2.5 = 1.9$ kc). These estimates are in good agreement with Figure 18 of Chapter 3.

Hydrophone directivity therefore confers two advantages: (1) discrimination against masking sound which come from directions outside the main lobe, although side lobes may sometimes be troublesome; and (2) angular discrimination, or ability to scan a restricted solid angle for a target. This first advantage is not directly relevant to masking criteria, although it is important in practical performance. The amount of this discrimination changes with frequency, and therefore alters the slope of the masking spectrum, but, within wide limits, the slope of a distributed spectrum has little influence on primaudibility (see Section 4.2.5). The second advantage may have a considerable effect on the recognition differential obtainable in practice. When a hydrophone with good angular discrimination in the search plane is trained across a suspected bearing at the proper rate, the resultant auditory motion should provide a helpful cue in the detection of signals, and should also increase the operator's confidence in reporting contacts and bearings.

Consider, for example, the situation in which (1) the primaudible target component is a sustained, unmodulated tone, (2) the background is ambient noise, and (3) the mean intensity and composition of the received background is essentially the same at all bearings. If the tone radiated by the target is primaudible at the listening hydrophone, a change in the quality of the received sound occurs when the hydrophone axis crosses the target. This change of quality will be detected only when the tone-to-noise ratio equals or exceeds the primaudible value, but the latter value appears to be several decibels lower for a modulated tone than for an unmodulated tone (see Section 4.2.2). By repeatedly sweeping the hydrophone axis across the target bearing, the operator can, in effect, introduce an amplitude modulation in the received signal, or rather, in the signal-background mixture, and thereby detect a fainter tone than would otherwise be possible.

The locally produced time pattern should be most effective (1) when the level of the mixture varies over a range of 1 decibel or more, since smaller changes are likely to escape detection, and (2) when the change is repeated at a rate of about 3 cycles, since slower changes place a greater burden on the memory, and faster ones become difficult to perceive. For optimal results, therefore, the intensity of the received mixture should pass smoothly from its lower to its upper limit in about 0.16 second, and perform the reverse change in a like time; in other words, the level of the mixture should rise or fall at a rate of about 7 decibels per second (1 decibel per 0.16 second). If practicable, the transit should be repeated several times at a regular rate of about 3 sweeps per second; therefore the angular limits of train should not greatly exceed those needed to drop the nonaxial intensity by more than the required 1 decibel.

Let I represent the average intensity of the received background, in the critical band centered at the frequency of the primaudible tonal component in the signal; i_p the peak intensity of the tone, received in the on-target position; and i some lower intensity of the tone, received when the hydrophone is trained ϕ degrees off target. From the preceding discussion, the hydrophone must be trained between limits such that $(I+i_p)/(I+i) = 1.25$, since 10 log $1.25 = 1$ db. This expression contains three unknowns; therefore one additional condition must be imposed before it is possible to obtain a relation between i and i_p, and thereby determine ϕ, the number of degrees the hydrophone must be trained off target during the process of sweeping. This additional condition is the value of i_p/I, the ratio at primaudibility between the recurrent peak intensity of a modulated tone and the intensity of the distributed background included within the critical band centered at the tone frequency. This value cannot be stated with certainty. From Table 4, in Chapter 4, a reasonable value appears to be $i_p/I = \frac{1}{2}$. For the sake of generality, however, two additional possibilities are considered: $i_p/I = 1$ and $i_p/I = \frac{1}{4}$. In decibels, these correspond, respectively, to a tone primaudible when its level is 0, 3, and 6 decibels below the level of background within the critical band. The values of i/i_p obtained from the first condition, upon substituting each of the three alternative values of the second, are listed in Table 1 in arithmetic and also in logarithmic form.

It is clear from the last column of Table 1 that even 100 per cent modulation is not likely to render the tone primaudible when its recurrent peak level is about 6 decibels below background within the critical band which includes the tone. Furthermore, improvement as great as this may be expected only if it is valid to assume that an increase of 1 decibel in a single critical band can be perceived, even though no corresponding changes occur in the other critical bands stimulated by background. Table 1 also implies that when sweeping a hydrophone across the target, the sweep should be great enough to reduce the target sounds by at least 3 decibels from the level received on the hydrophone axis (see bottom line of table).

Thus, for many practical purposes, it seems most suitable to adopt the angle between the axis and the −3-decibel point as numerically defining the *angular discrimination* of a hydrophone. This is generally about one-fourth the width of the main lobe, defined in Section 3.3 as the angle between the two directions of minimum response on each side of the hydrophone axis.

TABLE 1. Estimated effect of modulation on detection of a masked tone.

$\frac{i_p}{I}$	1	0.5	0.25
$10 \log \frac{i_p}{I}$ in db	0	−3	−6
$\frac{i}{i_p}$	0.6	0.4	0.0
$10 \log \frac{i}{i_p}$ in db	−2	−4	−∞

When marked modulation, or a similar time pattern in the incident signal, does not make the process superfluous or harmful, it would seem that sweeping a directional hydrophone through the target bearing at the optimal rate should improve the detectability of tones by 2 to 3 decibels. The value of this listening technique in practice may be better appraised if it is recalled that when signals are slowly faded into the presented sound, as in the 2-hour tests described in Section 4.2.6, the performance of observers becomes inconsistent and generally requires 2 to 3 decibels more gain in the signal channel to assure detection. It seems unlikely, however, that the combined improvement of 4 to 6 decibels above performance in the 2-hour tests would usually be obtained in the field because the noise background, in contrast to what was assumed in the preceding discussion, may itself show a disconcerting dependence on hydrophone orientation, and because, as indicated later, there is probably no unique optimal rate of train and no practicable method of maintaining such a rate if it could be specified. It seems likely, however, that the field *RD* may be taken as essentially equal to the random-order *RD* when the operator can train a directional hydrophone. It should be noted here that the on-off effect, in addition to improving responsiveness to a given signal component, may offset the effects of careless search or poor presentation, and thereby increase the chance that an operator will detect the optimal component (see Section 4.2.6). Familiar illustrations, from the field of vision, of the principle that intermittent stimuli may be more attention compelling than sustained stimuli of the same kind are the flashes received from lighthouse beacons and from illuminated advertising displays. A few of the tests described in Section 5.6 indicate that the ability to detect the optimal component is enhanced by sweep modulation.

To apply the preceding discussion, consider a JP-1 hydrophone; let the signal be a sustained 5-kilocycles tone, radiated by a distant target, and just audible in the presence of isotropic noise. The width of the main lobe is equal to $2\lambda/d$, which at 5 kilocycles gives 0.64 radian, or 37 degrees (compare with Figure 16 in Chapter 3). The angular discrimination (or approximate arc which must be described to produce a detectable change in the character of the mixture received in the on-target position) is one-quarter of the lobe width, or 9 degrees. Thus the total angular sweep is 18 degrees. For an assumed modulation rate of 3 per second, this corresponds to an average sweep rate of 54 degrees per second, or about 8 rpm (see Table 2). Under the same conditions, a 5-foot line would have to be trained at only 3/5 the rate of the JP, because the lobe width is inversely proportional to hydrophone length. The optimal rate of sweep also depends on the frequency of the primaudible signal, since lobe width is inversely proportional to frequency. From these relations, estimates may be made of the optimal rate of sweep, and also of the angular discrimination and bearing accuracy, defined as the rms bearing error. A group of such estimates, showing the dependence upon hydrophone length and listening frequency, are given in Table 2 and are in good agreement with observation. The values for a 1-foot hydrophone at a frequency of 1 kilocycle are probably overestimates, since even an inefficient baffle will produce a noticeable change in character for off-axis orientations exceeding 90 degrees; these values have therefore been set off in parenthesis.

It will be observed that all the quantities tabulated decrease with increasing frequency and increasing length. The assumptions made in computing bearing accuracy are discussed later in this section.

TABLE 2. Estimated optimal sweep rate, angular discrimination, and bearing accuracies of line hydrophones for tonal signals detected in the presence of isotropic noise. Values in parentheses are probably overestimates if the hydrophone is supplied with a baffle.

Hydrophone length in feet	Frequency of tonal signal in kc								
	1			5			10		
	Optimal rate in rpm	Angular discrimination in degrees	Bearing accuracy in degrees	Optimal rate in rpm	Angular discrimination in degrees	Bearing accuracy in degrees	Optimal rate in rpm	Angular discrimination in degrees	Bearing accuracy in degrees
1	(120)	(±135)	(±45)	24	±27	±9	12	±14	±5
3	40	± 45	±15	8	± 9	±3	4	± 5	±2
5	24	± 25	± 8	4.8	± 5	±2	2.4	± 3	±1
8	15	± 15	± 5	3	± 3	±1	1.5	± 2	±1

The total angle of sweep under optimum conditions is twice the angular discrimination, given in the second column under each frequency. In practice, mechanical difficulties of training the hydrophone may require slower rates of sweep than those shown, or, alternatively, a wider sweep, which also would lead to an effective reduction in the rate of modulation. In particular, for hand-trained JP gear the moment of inertia of the hydrophone would make it difficult to maintain an angular oscillation at the indicated rate over the required angular limits, that is, 3 sweeps per second. Figure 64 in Chapter 4 suggests, however, that the rate of modulation can be changed by a factor of 2 without changing the recognition differential by more than 1 decibel.

It may be inferred that short, mechanically trained lines are undesirable because of their broad response patterns and poor discrimination against isotropic noise, and because of the high rates of train usually needed. Extremely long, mechanically trained lines would give good noise discrimination and good angular discrimination, but would be mechanically cumbersome; large forces would be required to train them, since the moment of inertia increases as the cube of the length, while the angular sweep rate and resulting angular acceleration for best aural performance is inversely proportional to the first power of the length. Other mechanical problems also become serious for long line hydrophones. For example, relatively small amounts of pitch and roll may make it difficult to maintain contact with a target, or to scan a given bearing effectively, when the width of the main lobe is too narrow.

There is, however, no reason to believe that any single length is optimal, or that any particular ratio of train is best. Practical listening situations can vary over a large range, and primaudible signal frequencies will vary correspondingly. A few general and fairly rough rules can be suggested from estimates[12] of the probable frequency at which primaudibility will occur under different operating conditions, but the forces and rates of train which would be required evidently cannot be maintained during an entire watch while searching for contacts. There is no justification, under the circumstances, for optimistic views concerning the inherent superiority of recognition differentials which can be obtained in the field with the angular scanning technique, as against recognition differentials obtained in the laboratory by the random-order method. But there is little doubt that performance in the field will be improved by using a listening hydrophone with good angular discrimination and that less uncertainty will be experienced at first contact when selective, angular scanning is possible.

The estimated bearing errors listed in Table 2 have been computed by assuming (1) that errors (made in determining the midpoint between the angular limits of train) follow an approximately normal distribution; (2) that the mean value of the distribution, in the absence of systematic error, is indistinguishable from the true bearing; (3) that the number of reported sonar bearings which deviate from the true bearing by an amount exceeding the average angular limits of train is negligible; and (4) that the change in true bearing during the several successive transits required to obtain a sonar bearing may be ignored. It is a commonly used rule-of-thumb that practically all of a normally distributed population is included between limits which differ from the mean by 3σ, where σ is the standard or rms deviation. Accordingly, if the angular limit of train be equated to 3σ, the estimated standard deviation of the bearing errors is one-third of the average limit of train; this criterion was used in obtaining the bearing errors listed in the table. The rms deviation from the mean in a normal distribution includes $2/3$, or more exactly 68.3 per cent, of all the cases; in other words, if the preceding analysis is valid, two of every three bearing determinations will be in error, on the average, by no more than the tabulated values. While this argument is not precise, it yields conclusions in good agreement with rms bearing errors measured for a variety of conditions.[6,8]

Because of selective attenuation, the sonic detection of targets which become primaudible at ranges in excess of 7,000 to 10,000 yards tends to occur at frequencies below 2 kilocycles.[12] In this region, the required rates of train are high and the available bearing accuracies relatively poor (see Table 2). This conclusion agrees with the practical observation that the bearing accuracy of 3-foot lines deteriorates when the target vessel is detected at very long range, according to an informal communication from CUDWR-NLL. At 1 kilocycle, the discrimination of the JP hydrophone against deep-sea ambient is about 2 decibels, that of a 5-foot line is about 4 decibels, and that of an 8-foot line is about 6 decibels (see Figure 18 in Chapter 3, and the preceding discussion). These same relative differences apply to all frequencies between 0.5 and 10 kilocycles. It is possible that similar statements apply to self-noise, but there is no direct evidence on this point. The different amounts of discrimination against nondirectional background imply that detection ranges will be somewhat greater for the longer lines and that the primaudible frequency, i.e., signal-to-noise ratio in optimal band, may vary with the listening hydrophone for a given signal (see Table 3 in Chapter 4). When the frequency at which a given target becomes primaudible does depend on hydrophone length, the quantities listed in Table 2 should be compared along a diagonal rather than within a column.

When the signal-to-noise ratio is equally good over a fairly wide band of frequencies, training the hydrophone will change the received level of the higher frequencies more rapidly than that of the lower, because the width of the main lobe decreases with increasing frequency. If the signal-to-noise ratio is well above the value required for primaudibility, a change in quality will be heard as the hydrophone is trained through the target. However, at primaudibility it is likely that the better time pattern in the high-frequency portion of the optimal band will render it more audible than the low-frequency portion, and the effective band of primaudible frequencies will be narrower than might be expected on the basis of the signal-to-noise ratio alone.

Since most target spectra are distributed sounds, the changes in quality just mentioned may assist in obtaining sonar bearings when the signal is well above primaudibility. To take advantage of this factor, however, the operator must be able to concentrate on the high signal frequencies and disregard the low ones, since the target will be audible over a fairly wide band of frequencies and the intensities of the low-frequency components will change relatively slowly with angle of train. If the operator is unable to ignore the low-frequency components, bearing accuracy will suffer. This problem may be met by inserting the proper high-pass filters, if these are available, or by dropping the gain until the signal approaches the audibility threshold. The effect of the latter is equivalent to introducing a high-pass filter (see Figure 5 in Chapter 4).

The discussion of the present section is intended to serve as a basis for coordinating and assessing observations described in the remainder of this chapter.

5.2 TECHNIQUES

The listening tests[8,9,10] were conducted aboard a laboratory vessel anchored offshore. A signal of controlled intensity was radiated by an underwater projector mounted astern, and the hydrophone under test, suspended to the depth of the underwater projector through a well in the hull, picked up a mixture of signal and background. The latter consisted largely of water noise and sounds produced by the slapping of waves against the hull. Thus, the character of the mixture received at various orientations depended upon the properties of the hydrophone in use. Since the distance between the listening hydrophone and the artificial target was fixed, and the energy radiated by the target could be controlled, it was possible to evaluate recognition differentials under semifield conditions. However, variability of such factors as ambient noise level and transmission conditions, such as are usually found in the field, may have influenced the results.

The four signals used were originally obtained as high-quality disk recordings. The recording hydrophone was nondirectional and had uniform response over the sonic or recording band. The sounds radiated by the four target vessels were received in quiet water and at a range of about 200 yards. For the listening tests, a suitable portion of a disk recording was impressed on an endless magnetic tape capable of continuously repeating a signal 20 seconds long. The magnetic playback technique has the advantage of extending the useful life of the disk, and is reliable even when the listening-test vessel sways with large amplitude.

As used in these tests, the underwater projector may be considered a point source of sound, radiating equally well in all directions and at all frequencies between 0.1 and 10 kilocycles. The projector and listening hydrophone were 40 feet apart. Acoustic coupling effects between a broadly tuned source and receiver are probably negligible at this distance,[a] but the curvature of the wave front incident on a hydrophone so close to the source may produce significant variations in receiver response. The two circular arcs in Figure 1 represent

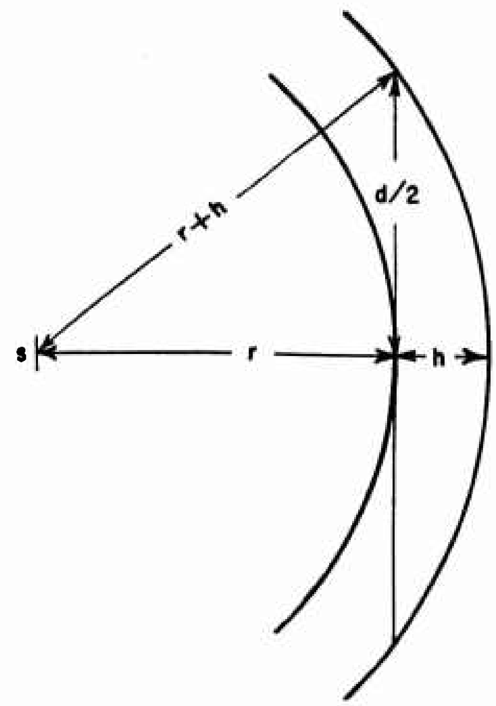


FIGURE 1. Spherical wave front incident on a hydrophone. In the plane of the figure, the receiver has a length d. The radial distance from the source s to the midpoint of the receiver is r, and to either end of the receiver is $r + h$. Hence, $(r+h)^2=r^2+d^2/4$; or, provided the term in h^2 may be neglected, $h=d^2/8r$. If $h=\lambda/4$, $f=2rc/d^2$, where λ is the wavelength, c the velocity, and f the frequency of the incident sound. For $r = 40$ feet, $d = 8$ feet, and $c = 4{,}800$ feet/second, $f = 2rc/d^2 = 6$ kilocycles and $\lambda << 4r$ (as assumed in dropping the term involving h^2).

parts of a progressive spherical wave which are separated by a distance h equal to one-quarter the wavelength. In this condition, the ends of the receiver lag its center by 90 de-

[a] Absence of electrical interaction between the signal and listening hydrophone circuits was demonstrated by surrounding the projector with several layers of Airfoam rubber, to reduce radiation into the water, and then driving the source at maximum electrical power. Under these conditions, no measurable or audible signals were observed at the outputs of the test hydrophones.

grees. For larger angles of lag the integrated output from the various segments of the hydrophone surface begins to drop sharply, due to mutual phase cancellations. It follows that, for frequencies exceeding $2rc/d^2$, the axial hydrophone response falls below the value obtained with a plane wave. For an 8-foot line-hydrophone, the longest used in these tests, the limiting frequency is 6 kilocycles; for a 4-foot line, good response is maintained up to 24 kilocycles. When curvature of the incident wave front impairs response, it also broadens the effective lobe width, since phase differences among segments of the receiving surface are changed less in orienting from on-axis to off-axis positions. It is evident that these effects resulting from curvature of the wave front were probably negligible in these tests, except possibly with the 8-foot line.

Measurements of the received signal, or listening hydrophone output, were made only when the axis of the test hydrophone was on target and the overall level of the received signal was well above overall background. The background present was also measured while the projector was silent. These measurements of received levels were made with a meter adjusted to follow peak amplitudes. For most sounds measured, the indications given by this meter fluctuated from the mean reading, which was used in computing recognition differentials, by about ± 2 decibels. The overall readings measured power in the 0.2- to 10-kilocycle band.

No accurate and independent measurement of the signal level received at primaudibility can be made under field conditions, because that part of the listening hydrophone output corresponding to signal is a small fraction of the total due to signal and background, and because fluctuations in the level of the listening hydrophone output usually exceed the contribution made by the signal. In order to evaluate field recognition differentials it is consequently necessary to know the relation between the level of the received signal and the known input to the projector. The recognition differentials given later in Figures 5 through 7 were obtained by assuming that the intensity of the received signal was directly proportional to the power dissipated in the projector circuit, in other words, by assuming that measurement of the power dissipated, at and above primaudibility, and of the power received, above primaudibility, permit calculation of the received primaudible signal power by simple proportion.

5.3 SIGNALS AND BACKGROUNDS

Four signals were used, and these showed various degrees of amplitude modulation. Three of the signals were recordings of submarine sounds; the fourth, a recording of sounds from a surface vessel. The character of the submarine sounds, at and above primaudibility, is described in the legend of Figure 5. It should be noted that these descriptions apply to a specific set of conditions—hydrophone, listening band, signal level, among others. The character of the ship signal, above primaudibility, is described in Section 5.6.

Frequency analyses were obtained for the three submarine signals. These were made by scanning the output of the test hydrophone, in the interval 0.2 to 10 kilocycles, with a 50-cycle band-pass filter of variable midfrequency. During such an analysis, the hydrophone axis was held on target, the received overall signal level was well above overall background, and the power dissipation in the projector circuit was maintained at a measured value. The levels in successive 50-cycle bands, relative to the overall level of the received sound, were obtained with a power level recorder adjusted to read rms amplitudes. When the sounds measured are of constant amplitude, the joint use of a peak indicator for overall measurements and an rms indicator for analysis involves no difficulties. But when the sounds fluctuate and the resolution time of the indicator affects the reading, the relation derived between signal and background at primaudibility may be different from that found by the methods of Section 4.2.2. Levels in the various 50-cycle bands were found to fluctuate by between ± 3 and ± 5 decibels from the mean levels which were used to plot the spectra. The magnitude of this variability and the observation that the power level recorder shows more fluctuation in a 50-cycle band than in a 10-kilocycle band agree with the data described in Sections 4.2.2 and 4.2.3.

The spectra in Figures 5 through 7 were transcribed from the power level recorder trace by reading the relative levels at 100-cycle intervals below 1 kilocycle, and at either 500-cycle or 1,000-cycle intervals above 1 kilocycle; levels at intermediate frequencies were plotted only when they deviated markedly from the general trend. Successive points in the plotted spectra were connected by straight lines. This method of representing the 50-cycle analysis lacks the clarity and precision of the point-by-point procedure used in preparing the spectra shown in Section 4.2.2, but when the depicted sound resembles cavitation or ambient noise, the two methods should yield essentially the same trend.

The spectra in Figures 5 through 7 represent the hydrophone output and do not exhibit the characteristic negative slope shown by the spectra illustrated in Sections 4.1 and 4.2. The latter were obtained with recording hydrophones which had an approximately uniform response over the recording band. The projected signals used in the New London tests also had the typical negative slope of ship sounds, but receiver response increased with frequency and thus offset the falling characteristic of the sound-in-the-water. Figure 2, for

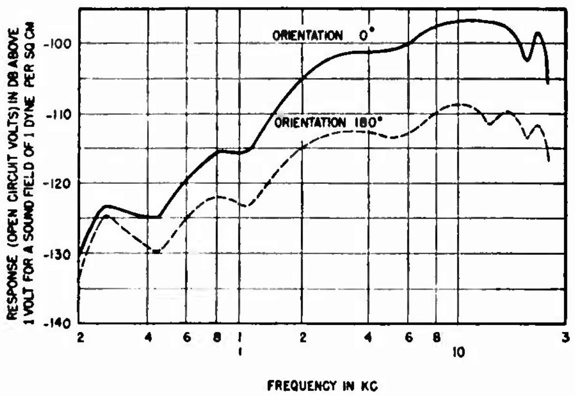


FIGURE 2. Frequency response of the JP-1 hydrophone-baffle assembly in a plane-wave field.

example, shows the frequency response of a JP hydrophone. The ordinate gives response in terms of output voltage generated by a plane wave, incident along the acoustic axis, which has an rms pressure of 1 dyne per square centimeter at each calibrating frequency. The indicated response characteristic rises at a rate of about 9 decibels per octave and drops abruptly at frequencies above 20 kilocycles.

This behavior is due to the combined effect of several factors, of which two will be mentioned here. To begin with, the voltage generated by variations of magnetic flux is proportional to dp/dt, the time rate of change of the incident pressure, and dp/dt, in turn, is proportional to the incident frequency. This factor alone would cause output to increase at a rate of 6 decibels per octave, since the power is proportional to the square of the generated voltage and hence to the square of the incident frequency. The drop above 20 kilocycles is due to the fact that the 2-inch diameter of the magnetostrictive cylinder is less than sonic wavelengths; thus, diffraction occurs and there is no well-defined acoustic shadow at the rear of the cylinder. Near 20 kilocycles the pressure variations, at the portions of the cylinder surface facing toward and away from the source, are 180 degrees out of phase; in other words, the diameter of the cylinder equals a half wavelength. Therefore the integrated response falls.

Practical listening installations whose response increases with frequency tend to compensate for the negative slope of ship signals and to present a flat spectrum to the headphones, as shown in Figures 5 through 7. Thus the relative loudness contribution from the middle frequencies is raised, and it becomes difficult to increase the sensation levels at the low and high frequencies without producing discomfort (see Section 4.2.5).

The masking background used in these tests was thoroughly realistic but varied from time to time in composition and intensity. In order to obtain reliable comparisons between hydrophones or observers under these circumstances, it is necessary either to monitor the background constantly, to conduct a group of simultaneous tests with one target and several hydrophones, or to run tests on different hydrophones in rapid succession. Thus, field studies are inherently more complex than laboratory tests and may yield less precise results.

The major sources of background were whitecaps, waves breaking against the hull, distant surf and harbor traffic, activities on board (including training of the test hydrophone), and system noise. Contributions from water motion depend on wind and sea states and on the orientation at which the listening vessel receives the impact of waves. Turbulence about a rapidly trained hydrophone, as well as mechanical vibrations associated with the process of training, undoubtedly contribute to the received background. This factor was not taken into account in measuring background levels, probably because such effects were too small to attract attention and hence too small to affect the results.

Figures 5 through 7 indicate that the composition as well as the intensity of received background changed with sea state, and that noise discrimination is better with long hydrophones and at high frequencies. The measured level of overall background received by a given hydrophone was 8 to 10 decibels higher for sea state 3 than for sea state 1; this increased contribution was restricted to the low and middle frequencies, with an actual decrease occurring in the high-frequency region. The report[8] describing these tests discusses the change of background with sea state as follows: "This type of change in the background spectrum between sea states 1 and 3 was caused chiefly by the great increase in the size and number of waves slapping the hull of the listening vessel, thus giving a rise in low and middle frequency levels; but at the same time, the number of whitecaps in the sea state 3 was not quite as great as the number present in the sea state 1 because of differing wind conditions, and as a result high frequency levels were lower for sea state 3." This type of background behavior is somewhat unusual since these particular tests were conducted in a cove. The shoreline shielded the water from the full impact of the wind in the state 3 tests, but not in the state 1. Hence, the wave height, but not the number of whitecaps, corresponded to a number 3 sea in open water. It is possible, however, that the effect is general and that with increasing sea state, the effect of wave slap on the hull may increase the received noise background at low frequencies more than wave slap, whitecaps, and other factors increase the high-frequency noise components (see also Figure 4 in Chapter 4).

Figure 3 shows nearly simultaneous power lever recorder traces of background received in the 0.2- to 10-kilocycle band by several hydrophones with different directivity indices. Reference 8 discusses these traces as follows: "In sea state 3 there were many localized sources of noise arising from waves breaking near the listening vessel and wave slap on the hull. The random occurrence of these noise sources gave integrated 'smooth' character to the sound received in the nondirectional hydrophones. However, when hydrophone directivity was introduced the number of sources affecting the character of sound was reduced and the integration

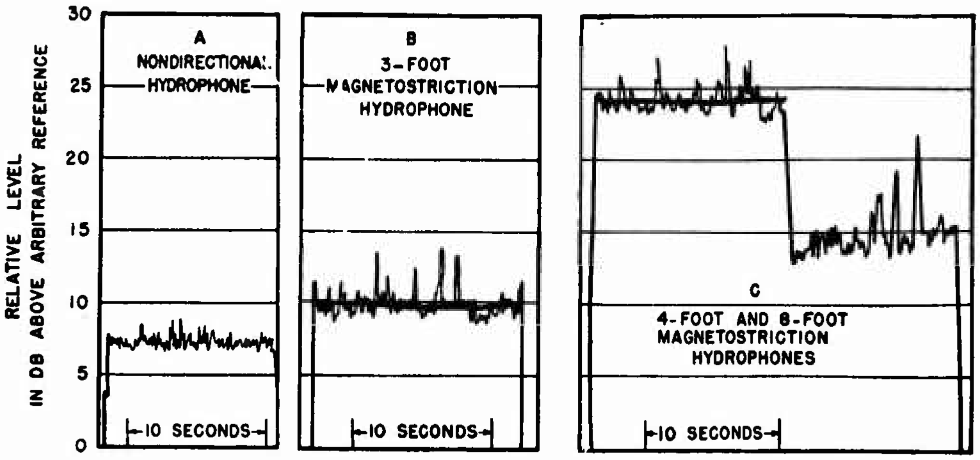


FIGURE 3. Time-amplitude pattern of water noise received by directional and nondirectional hydrophones. Mean levels of the traces are arbitrary and are not responsible for differences in the relative amplitudes of noise peaks.

of the sounds was not as complete; so that the individual sources began to become more distinguishable as individual splashes. Hence, the greater the directivity became the more important was each source of noise near the main lobe of the hydrophone. As a result the amplitude-time pattern of the directional hydrophones had much greater peaks and valleys than the nondirectional hydrophones. This is illustrated in Figure [3]. During the tests, comments of the listeners indicated that when listening with directional hydrophones their attention was focused only on the valleys of the amplitude-time pattern of the background noise and the peaks were ignored. Hence, the effective masking level of the noise was less than the average level; and since the volume indicator used in the tests responded to values between peak and rms readings, the calculated values of *RD* would be greater for patterns such as those in Figure [3C] than in Figure [3A]." The effect of "plonks" in received water noise is also illustrated in Figure 4 of Chapter 4; in that case also the effect of wave slap raised the low-frequency and middle-frequency levels to a greater extent than those at the high frequencies. It will also be noted that the simulation of hydrophone directivity described in Section 4.1.8 made no allowance for changes in the peakiness of background received with directional hydrophone.

The horizontal lines superposed on the traces in Figure 3 indicate how levels of masking background may be read from such traces in order to compute recognition differentials and to obtain the relation between signal and background spectra at primaudibility. It is probably justifiable to disregard the possibility of persistent masking (see Section 9.2.2), when the successive noise peaks occur several seconds apart, as in the case illustrated; thus, it is doubtful whether limiters would be of much help under these conditions. It might be useful in further recognition studies (in which the masking background fluctuates widely and so slowly that resolution time of the indicator does not falsify results) to examine the time pattern in the primaudible, as well as in the overall band. It is also worth inquiring whether or not increased peakiness in the background received by directional hydrophones increases the difficulty of maintaining faint contacts and getting accurate sonar bearings.

5.4 TEST METHODS

About a dozen different hydrophones were studied in this test program. They were usually supplied with a baffle and were coupled to the listening amplifier through one of a set of impedance-matching transformers. The performance of hydrophones is determined by a complex set of factors. In addition to those already discussed, such as angular and noise discrimination, there should be mentioned frequency response, listening band width, impedance match to amplifier, distortion, absolute response, and system noise. If the installation is unsatisfactory in any of these respects, signal detection may deteriorate. Little detailed information about these hydrophones, along the lines just specified, is currently available; hence, observations on comparative performance must be treated with caution in so far as the general problem of detection is concerned. It should be pointed out, however, that the major function of these comparative hydrophone tests was to select from among available designs the one best adapted to serve a specific purpose.

The nominal listening band used in this study extended from 0.2 to 10 kilocycles, but in a number of the tests high-pass filters were inserted in the amplifier circuit. The approximate values of the low-frequency cutoffs, as well as the effect of the filters in modifying the spectra of received signal and background, are shown in Figure 6. The characteristics, including insertion loss, of a similar filter set, built into the JP-1 amplifier unit, are given in Figure 4. The "bass-boost filter" indicated in Figure 4 is essentially a broadly tuned resonance network capable of accentuating the low-frequency range without unduly attenuating the high.

The effective listening band width is also determined by the properties of the headset. Two types of headsets, a crystal and a magnetic, were used during the course of these tests. While both types of phones were of high quality, their response characteristics[9] dropped sharply above 7 kilocycles.

Amplifier gain was set for a comfortable listening level. Seven different presentation

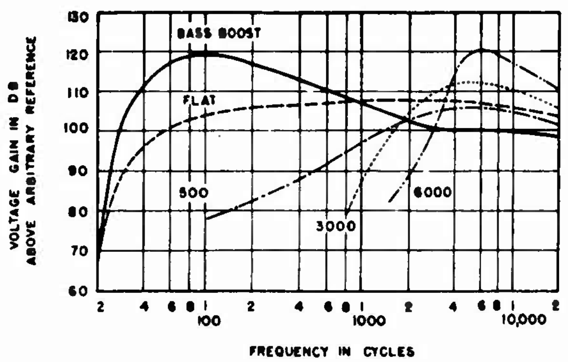


FIGURE 4. Response of the JP-1 amplifier for various high-pass filter positions. Numerals designate nominal low-frequency cutoff in cycles.

methods were tried at various times, but only one of these was used in a given test. The purpose of varying presentation method was to evaluate the effects upon signal detection of auditory motion introduced by training a hydrophone and by opening or closing the range between the target and listening vessel. The two extremes of auditory motion among the seven methods examined correspond to the sustained and the interrupted signals used in the laboratory tests described in Section 4.2.6; in other words, the hydrophone axis was fixed on the projector bearing, the gain in the projector circuit was maintained at a selected level, and a disabling switch made it possible to radiate either a sustained signal or one which was completely blanked for brief intervals. Clearly, the use of a blanking switch makes it possible to get much sharper off-on effects than can be approached in practical listening unless extremely directional hydrophones are used. While care is necessary to eliminate false cues due to key clicks and switching transients, the chief effect of these disturbances is usually to scatter signal energy beyond the limits of the frequency band containing the uninterrupted signal (see Section 7.1); hence, transients are not likely to falsify results when the masking background extends over the entire sonic listening band.

To test the effects of increase or decrease in target range when the listening hydrophone is not trained or has poor angular discrimination, the axis of the test hydrophone was fixed on target and the level of the radiated signal raised or dropped at the rate of 0.2 decibel per second. This is a rather low rate of auditory motion but probably defines the upper limit of rates likely to be encountered in practical listening with a fixed or nondirectional hydrophone. Even at a range as short as 1,000 yards and a speed of about 10 yards per second (18 knots), the change of average intensity amounts to only 0.09 decibel per second, provided the intensity is assumed to fall off as the square of the distance. In the 50 seconds required to close the range from 1,000 yards to 500 yards, on these assumptions, the sound level will increase by only 6 decibels, giving an average change of 0.12 decibel per second during this interval. At longer ranges or with slower ship speed, even smaller rates of change will be found. It should be noted, however, that much larger intensity rates may be produced occasionally by factors causing rapid fluctuations in transmission (see Figure 13 in Chapter 3), but such fluctuations are irregular and may cause the signal-to-noise ratio to fall below the primaudible value and thus hamper detection.

Three additional presentation methods were devised in order to study the effect of joint motion of the hydrophone and the target. In each case, the test hydrophone was trained through the projector bearing at about the optimal speed, and the level of the radiated signal was held at a selected value or changed progressively at a rate of 0.2 decibel per second. It is possible that such effects could be simulated adequately in the laboratory by using the proper combination of filter elements and time constants in the signal channel, and synchronizing their action with the motion of a bearing wheel controlled by the subject.

To determine primaudible signal levels for the various presentation methods, the gain setting in the projector circuit was changed and a tabulation made of power level and correlated listener response. Primaudible components were identified solely on the basis of the observers' subjective impressions.

When several listeners were available simultaneously, a test consisted of a group of 20 to 30 presentations, each lasting about 10 seconds, and the test administrator varied successive signal levels in random fashion. When the sound level was constant during presentation, the observers removed their phones between presentations, in other words, they did not listen while the signal was being adjusted to its new level. When signal level changed continuously during the listening interval, each observer noted the time at which the signal became audible or inaudible. From the recorded times, the mean result for the group could be determined. During group tests, the hydrophone was controlled by the test administrator, who judged the proper rate and limits of train by aural monitoring. In practical listening, the coordinated activity of an operator's arm and ear may assist signal detection; conversely, the observers in these tests had the advantage of knowing that a signal was usually present in the received mixture, although occasional checks were run with the projector disabled.

A large fraction of the recognition differentials were obtained in self-administered tests performed by experienced listeners, who started with an inaudible signal level and increased it in steps of 2 decibels until the masked threshold was reached. In some cases, the rate and direction of change was determined by the presentation method selected. Self-testing was preferred for its brevity and also made it possible for the listener to train the hydrophone himself.

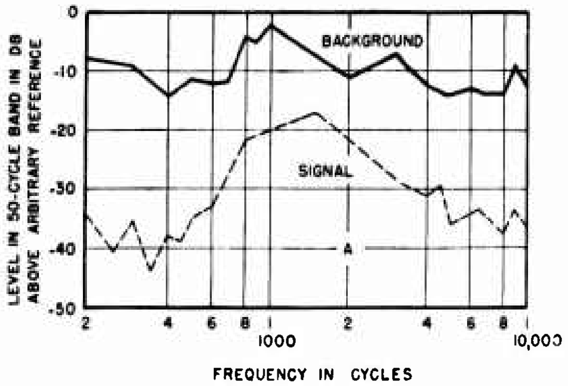
The inconstancy of background made it difficult to correlate the various methods of test administration and scoring. A few comparison checks were made, however, and the observed variation among the averages obtained for any two techniques or groups of listeners did not appear to exceed 1 to 2 decibels. Mean results, based on groups of between 2 and 8 observers, were obtained under as nearly constant background conditions as possible, and the index sought was that signal-to-noise ratio at which only half the group could hear the signal. No transition curves are available for these tests. However, in Chapter 4, Table 6, and Figures 6 through 23 imply that the slopes of such curves are not strongly dependent on signal modulation, and thus that the recognition probability curves obtained without benefit of hydrophone sweep are applicable to the field situation.

5.5 OBSERVED RECOGNITION DIFFERENTIALS

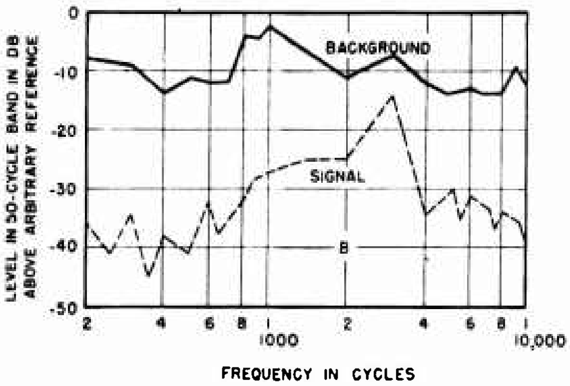
Allowing for the difficulties and the differences in procedure, the relations between primaudible signal and background spectra deduced from these tests seem to be in substantial agreement with the trends discussed in Sections 4.1 and 4.2. Three general types of primaudible component were detected: tones, propeller sounds, and distributed noises with little distinctive character. The results obtained for each type are discussed in this section. Some observations concerning the effects of hydrophone characteristics and filters are also described in this connection.

Tones primaudible at 1.5 and 3 kilocycles are represented in Figures 5 and 6. Since the critical bands at these frequencies are about as wide as the 50-cycle analyzing band, detection of sustained tones would be expected when the levels of primaudible signal and background spectra are very nearly equal (see Figures 76 through 79 in Chapter 4). Instead, the relation between the spectra indicated in five typical cases (Figures 5A and 5B, and 6A. 6B, and 6C) corresponds on the average to a signal-to-noise ratio of -5 decibels in the optimal band. As pointed out later, these cases represent a fairly wide sampling of test conditions; hence, the mean deviation from the average ratio is fairly large (3 decibels).

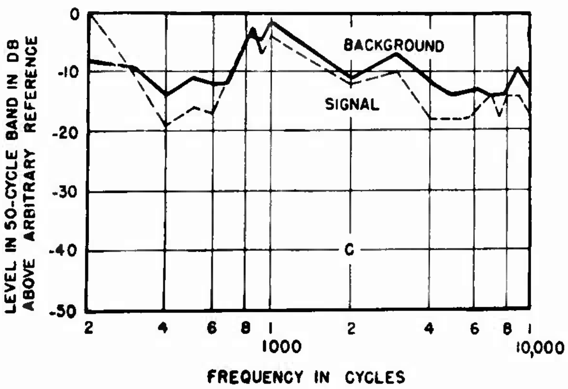
The average spacing of 5 decibels between signal and background spectra at their points of closest approach is probably due largely to the observed fact that the tones fluctuated by between ± 3 and ± 5 decibels from the mean levels shown in the figures (see Section 4.2.2) and that the primaudible signals were detected at their peak, rather than their mean levels. In addition, the individual spectra show the



A. Signal was sonic noise from an S-class submarine proceeding at 240 rpm (5 knots) and periscope depth, consisting of cavitation noise with moderate propeller thrash and also fluctuating tones at about 500 and 1,500 cycles. *RD* for presentation band shown: 13 decibels. Character of primaudible signal: 1,500-cycle tone and propeller cavitation in middle frequency region.



B. Signal was sonic noise from a fleet-type submarine proceeding at 100 rpm (5 knots) and periscope depth, consisting of cavitation noise with weak propeller thrash and also a strong sustained whine between 1 and 3 kilocycles. *RD* for presentation band shown: 14 decibels. Character of primaudible signal: whine and hiss.



C. Signal was sonic noise from a surfaced fleet-type submarine proceeding at 170 rpm (12 knots), consisting of cavitation noise with definite propeller thrash and also low-frequency grinding sound and irregular thuds and clanks. *RD* for presentation band shown: 1 decibel. Character of primaudible signal: grinding sound and fluctuating broad-band hiss.

FIGURE 5. Effect of signal spectrum on audibility. Background was water noise and wave slap, with sea state 1. The listening band extended from 0.2 to 10 kilocycles; the hydrophone used was a 3-foot crystal line. The presentation method consisted of interrupting the signal while the hydrophone axis remained fixed on the projector bearing.

influence of presentation method. Thus, Figure 5A, obtained by blinking the signal, shows a spacing of about 10 decibels at 1,500 cycles, while Figure 6, obtained by training the hydrophone, but under conditions otherwise similar to those indicated in Figure 5A, shows a spacing of about 6 decibels at the same frequency. In other words, square-wave modulation of the signal improved recognition by about 4 decibels. It should be noted, however, that auditory motion was present under both conditions of test; it was merely more drastic in one case than in the other. Effects produced by auditory motion are also discussed in the following section.

The spacing of 7 decibels at 3 kilocycles shown in Figure 5B is in general agreement

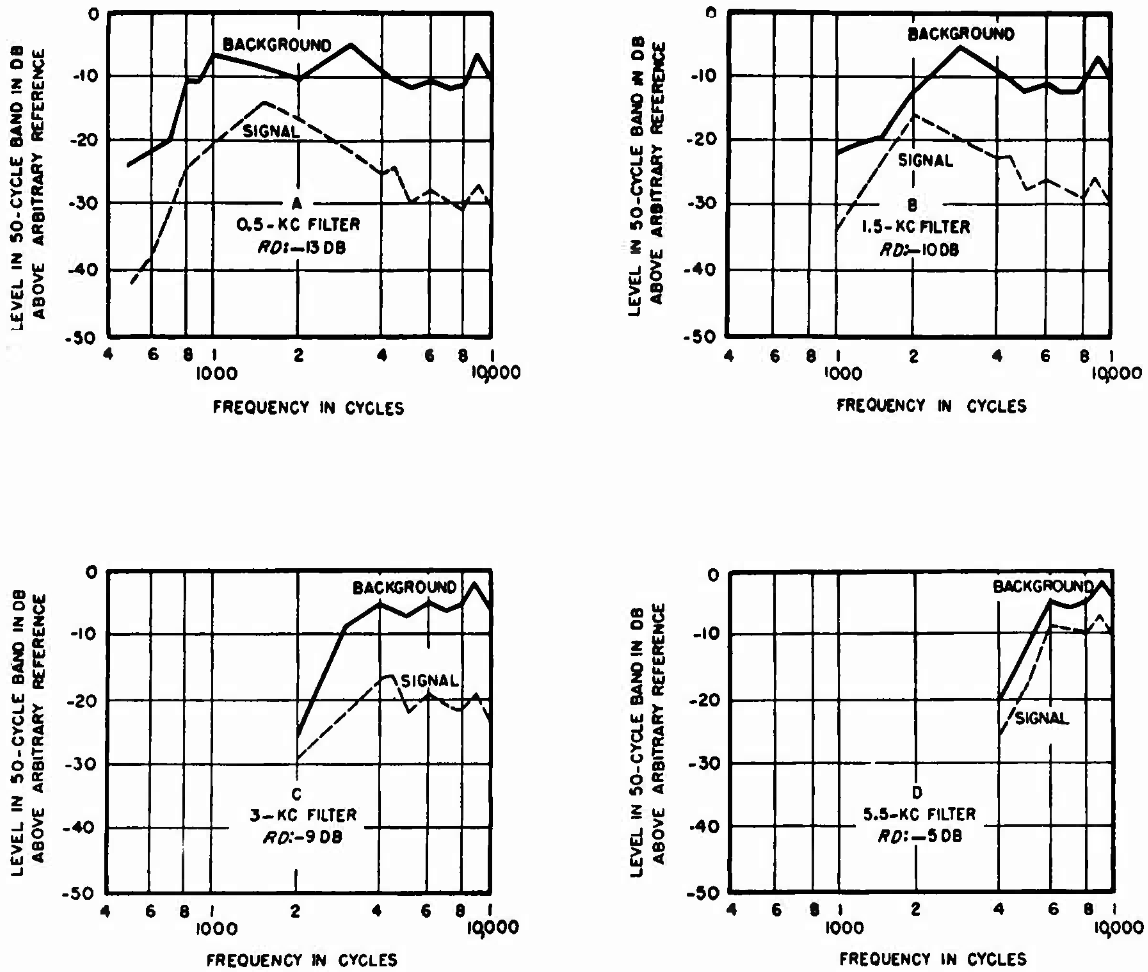


FIGURE 6. Audibility of sonic noise from an S-class submarine proceeding at 240 rpm (5 knots) and periscope depth, masked by water noise and wave slap (sea state 1). The hydrophone used was a 3-foot crystal line. The presentation method consisted of training the hydrophone through the projector bearing while the level of the projected signal was maintained constant; the listening band was restricted, as shown in the figures, by inserting high-pass filters. The recognition differentials refer to the indicated presentation band. Character of primaudible signal: propeller sounds, when 5.5-kilocycle high-pass filter was inserted; otherwise, 1.5-kilocycle tone.

with the preceding discussion. The spacing at 1.5 kilocycles indicated in Figures 6B and 6C (about 4 decibels and 0 decibel, respectively) is smaller than the spacing shown in Figure 6A. The character of the sounds which were presented in the tests illustrated in Figures 6B and 6C differed from those illustrated in Figure 6A in that insertion of the filters produced a reduction in the relative level of the primaudible component and possibly some distortion in the optimal frequency region, as well as limiting of the rate of sweep modulation by the time constant of the filter. Such effects would be expected to make detection more difficult.

It should be noted at this point that the signals shown in Figures 5 and 6 were radiated in such a way as to emphasize their low-frequency content. They are, therefore, not directly comparable with the signals shown in Figure 7, which were radiated with uniform frequency weighting and which consequently had the characteristic slope (in the water) of about −6 decibels per octave. This partially explains the observation, indicated in Figures 5A and 7, that a different signal component became primaudible in two sets of tests in which the same signal recording was used. In addition, the hydrophones indicated in Figure 7 have greater discrimination against high-fre-

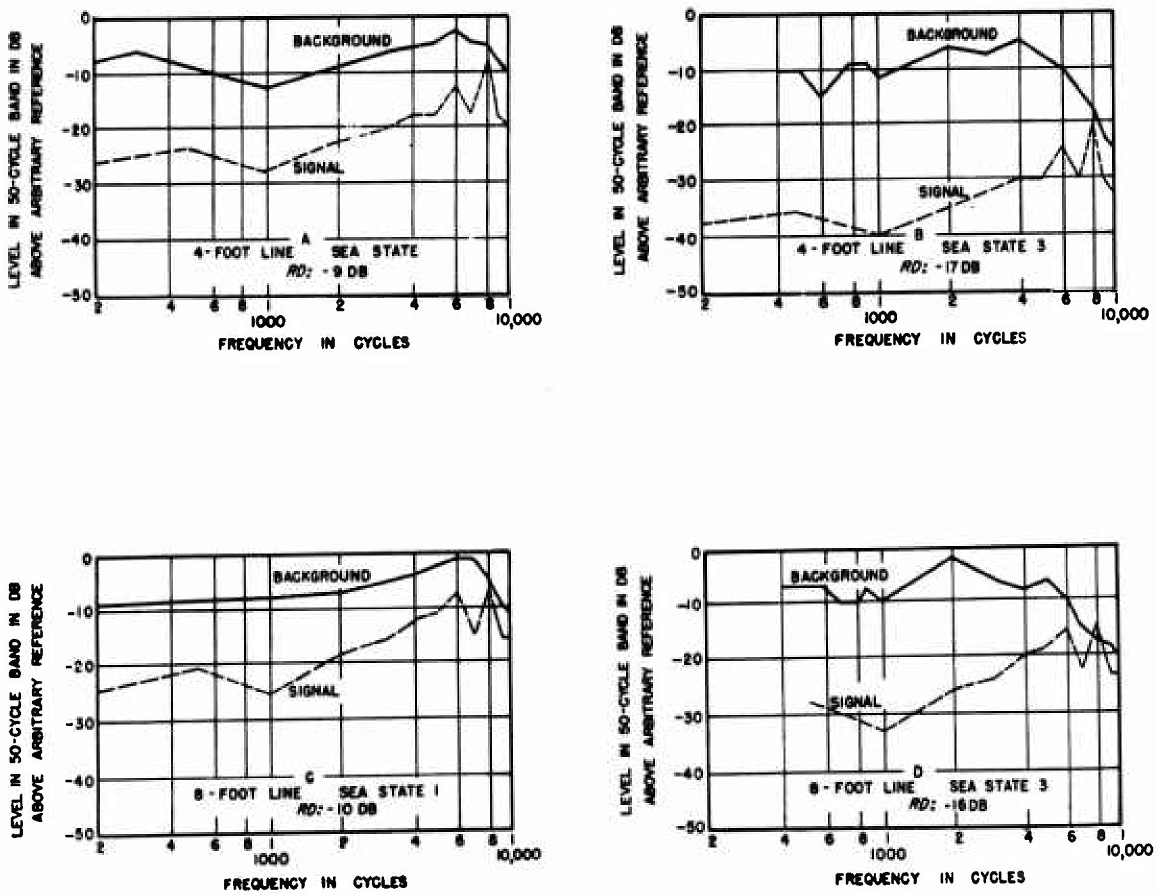


FIGURE 7. Audibility of sonic noise from an S-class submarine proceeding at 240 rpm (5 knots) and periscope depth, masked by water noise and wave slap (sea states 1 and 3). The hydrophones used were 4- and 8-foot magnetostriction lines. Presentation method consisted of training hydrophone through projector bearing while the level of the projected signal was maintained constant; the listening band extended from 0.2 to 10 kilocycles, and recognition differentials apply to this band. Character of primaudible signal: propeller sounds.

quency background noise than the one described in Figure 5; this factor, also, tended to shift the frequency of the optimal component.

The primaudible sounds indicated in Figure 5C were distributed rather than tonal. In this case, recognition occurred when signal and background spectra were essentially coincident throughout the presentation band, despite the square-wave modulation imposed on the signal. In fact, the signal-to-noise ratio in the band between 200 and 300 cycles is about 7 decibels. It should be noted, however, that the slopes of the spectra presented to the ear, as indicated in Figures 5 and 7, had a general trend of about 0 decibel per octave over most of the listening band. With this type of presentation, the frequencies near 200 cycles are usually threshold-limited rather than masking-limited when the gain is set for a comfortable loudness level (see Figure 79 in Chapter 4). Thus, the *RD* shown in Figure 5C could probably have been reduced by increasing the overall gain. For the same reason, 200 cycles is specified as the lower limit of the listening band used in these tests.

When propeller sounds were detected, the signal-to-noise ratio in the optimal band averaged -2 decibels for the five cases shown in Figures 6D and 7, and the mean deviation from this average was 2 decibels. These values apply to detection of only one signal component, which appears to have become primaudible as a band of frequencies heard in the neighborhood of 8 to 9 kilocycles, and which was described by the observers as a series of pulses of very high pitch. This component may represent a bearing squeak; in any case, the spectra in Figures 6D and 7 seem definitely to indicate that a marked prominence appeared in the high-frequency region of the radiated signal, and may imply that such peaks occur in the sonic output of ships and submarines (see also Figures 29 and 38 in Chapter 4). That this prominence between 8 and 9 kilocycles was present in the underwater signal used in these tests and not due to a peak in the response of the test hydrophones is indicated by the fact that it recurs in the spectra obtained with three different test hydrophones, and further by the fact that it is absent from the background spectra received with two of those hydrophones.

The available spectra do not indicate whether the strong component at about 8.5 kilocycles was a tonal or a distributed sound, but the fact that the signal-to-noise ratio in the optimal region was approximately 0 decibel at primaudibility implies that this component was probably distributed over a band about 500 cycles wide, which is the width of a critical band at 8.5 kilocycles (see the discussion of Figures 29 and 38 in Chapter 4). It will be noted too that, within the optimal band, the signal-to-noise ratio required for primaudibility of this component was smallest (most favorable to detection) for the least directional of the three test hydrophones (Figure 6D) and largest for the most directional (compare Figures 7A and 7C, or 7B and 7D). Examination of the data suggests that this deterioration may be associated with the increased peakiness of background encountered in tests with the more directional receivers (see Figure 3). Similarly a higher signal-to-noise ratio was required with the higher sea state (compare Figures 7A and 7B, or 7C and 7D).

One further precaution must be kept in mind when evaluating these tests in terms of practical hydrophone performance. When the level and character of background are reasonably constant, such comparisons between the performance of test hydrophones can be made by noting the power which must be dissipated in the projector circuit in order to assure detection with each hydrophone in turn. The overall signal-to-noise ratios measured at the outputs of the various hydrophones are not themselves an adequate index of comparative hydrophone performance since the level of the received noise may be considerably modified by the overall noise discrimination of a receiver, and there may nevertheless be only a very small effect on the level of the primaudible signal if the latter is detected at a frequency where hydrophone directivity is poor. In the field, such a hydrophone would give about the same performance as a completely nondirectional receiver, although the recognition differentials measured at the hydrophone outputs would be quite different in the two cases. It follows that the recognition differentials shown in Figures 5 through 7 cannot be used directly for evaluating practical performance.

The preceding analysis indicates that a variety of factors, such as discrimination against noise, auditory motion, ease of train, coupling, system noise and band width, time patterns in the received signal and background, frequency of the primaudible component, and occasionally others, may be expected to affect the minimum signal level detectable with a given hydrophone in a particular situation. A number of tests conducted during the program under discussion clearly reveals the desirability of using hydrophones with a large degree of discrimination against isotropic noise. However, since no adequate study of the effects of these factors could be undertaken during this test program, it is impossible to give any more detailed conclusions concerning comparative performance or optimal design characteristics on the basis of these findings. An examination of the effects to be expected from auditory motion is given in the next section.

It should be mentioned in conclusion that the filter tests described in the present study confirm (for a quasi-field situation, and for sounds presented to the ear with a slope of approximately 0 decibel per octave) several of the observations described in Section 4.1.8. Thus, it was found that narrowing the listening band by insertion of high-pass filters never made it possible to detect a given signal at a lower level. However, when the optimal component of that signal was not contained in the excluded portion of the spectrum, the use of filters made little difference. Occasionally, too, the primaudible component seems to have been detected below the nominal low-frequency cutoff of the filter (see Figure 6C), although at a less favorable signal-to-noise ratio. Finally, some observations should be mentioned which were made on an occasion when successive rain squalls and heavy seas briefly and repeatedly caused the level of background to rise and fall by 10 to 15 decibels. In this case, the observers did not themselves control the listening hydrophone, and use of a high-pass filter eliminated the more characteristic features of the ship signal. It was found that the observers could not detect the signal even when its spectrum was some 15 decibels above noise over the entire presentation band. They did hear changes of loudness in successive presentations but could attach no significance to them, owing to the absence of signal character and the high variability of background noise level.

5.6 AUDITORY MOTION

When a hydrophone can be trained back and forth across the target, the change in signal level constitutes auditory motion and may facilitate recognition. Comparisons based on tests made under fairly similar background conditions indicate the general importance of this effect. The reference datum in these comparisons is the masked threshold found when both hydrophone and mean signal level were held fixed.

"Blinking" the signal, while the hydrophone was stationary, helped most. The magnitude of this effect, which has no counterpart in practical listening, does not appear to exceed 4 to 5 decibels when the observers detect the same component in both conditions of test. About an equal degree of improvement was found when a highly directional hydrophone (4-foot or 8-foot line) was trained through the projector bearing at the optimal rate while mean signal level was held fixed; less improvement resulted when a hydrophone with poor directivity was trained through the target. The other presentation methods gave intermediate degrees of improvement. Signals of gradually changing level could be detected 1 to 2 decibels further below noise when they were initially audible than when their intensities were raised from below the primaudible level (see Section 4.2.6). Furthermore, this appeared to be true with trained as well as fixed hydrophones. Changes in the quality of the received signal-background mixture (associated with motion of the hydrophone) were easily heard when the signal was above its primaudible level, but such changes were almost imperceptible at primaudibility.[10]

Since the presentation and scoring techniques used in these tests are not strictly comparable with those described in Section 4.2.1, it is impossible to determine whether the reference datum mentioned above corresponds exactly to a 50 per cent *RD* as determined by the random-order method. If the shape of the transition curve changes markedly as a result of auditory motion, the change in *RD* for 95 per cent recognition, for example, may be quite different from the change in the 50 per cent *RD*. It seems unlikely, however, that this datum involves a larger signal-to-noise ratio than is needed for 80 per cent detection probability in the laboratory. Since the difference between the 50 per cent and 80 per cent *RD* will not change by more than 1 to 2 decibels with changing conditions, it may be inferred that the net improvement due to optimal auditory motion is probably not less than 2 to 3 decibels (see also Table 1).

It should be noted that all the results quoted above apply to the situation in which the same component was detected in the modulated and unmodulated signals. In at least two cases, however, it seems that a relatively strong component, which the observers failed to detect in the unmodulated signals, became the primaudible cue when modulation was introduced. In each of these cases, the component detected in the modulated sound permitted recognition of a fainter signal, and the recognition differentials were diminished by 10 to 12 decibels. The pertinent data will be found in reference 8.[b]

Similar observations are described in Section 4.2.2; listeners who failed to detect the optimal signal component required 6 to 8 decibels more gain in the signal channel than required by comparable observers. Thus, hydrophone sweep may compensate for the effects of inexperience or inattentiveness and, by directing attention to the optimal component, may occasionally double or triple the range at which a given operator can detect a given target.

All the observations discussed down to this point in the present section were made with two signals of rather different character. One of these was recorded from a submerged S-class submarine and became primaudible as a fluctuating 1,500-cycle tone (see Figure 5A). The other was recorded from a freighter with four-bladed screws driven at 60 rpm; it contained a great deal of clanking sound and had a heavily accented propeller beat followed by three minor swishes. No analysis was made of the freighter signal, but comments of the observers indicated that the primaudible component had a frequency in the region of 1 to 2 kilocycles. Since the degrees and rates of modulation of the optimal components in these signals are unknown, it is impossible to decide to what extent the improvement due to training the hydrophone is limited by modulations already existing in the radiated signal.

These observations appear to warrant the view that ability to train a directional hydrophone will improve performance under field conditions. However, the major advantage would probably be reduced variability of the operator's response (see Section 4.2.6).

[b] The effect of blinking the signal, using a relatively nondirectional hydrophone, is indicated in Table IVa, line 1, of that report; similarly, the effect of hydrophone sweep, using directional and nondirectional hydrophones, is depicted on the right-hand side of Figure 3 in the original report. See also page 17 and Table XII in that document. The second of these cases involved a change of hydrophone as well as of presentation method; hence differences in system response or inherent noise may have played a larger part than did the presentation method.

Chapter 6

SUMMARY OF LISTENING STUDIES

6.1 BASIC FACTORS

A PARTICULAR UNDERWATER sound or signal can be heard only if it is sufficiently intense compared with the background of unwanted sounds which tend to mask the signal. The signal level which permits recognition in half the trials is called the *signal recognition level*, measured in decibels. The difference between the signal recognition level and the background level is called the *recognition differential*. Since an increase in background noise requires a roughly similar increase in the signal recognition level, the recognition differential *RD* changes only slightly as the overall level of the presented sounds changes. The *RD* depends not only on the nature of the signal and background presented to the ear, but also on the method used for measuring signal and background levels. In particular the band widths used for such measurements must always be specified.

6.1.1 Signal Characteristics

The following characteristics of target sounds are significant in listening detection.

SPECTRUM

The measured sound power per cycle may show sharp peaks at certain frequencies, as a result of pure tones produced by vibrations of machinery on board ships. The sound from propeller cavitation shows a smooth spectrum, with the sound power per cycle present in the water dropping off by about 6 decibels every time the frequency is doubled. At considerable ranges the spectrum is modified by the change of transmission loss with frequency. Different types of targets may show characteristically different spectra.

MODULATION

Machinery sounds may show rhythmic oscillation. Propeller sounds are generally modulated at the *shaft rate*, the number of revolutions per minute (rpm); at the blade rate, which equals the shaft rate times number of blades on each propeller; or both. If directional listening gear is used to sweep across the target, modulation can be imposed on an otherwise steady sound.

6.1.2 Background Characteristics

Airborne noise may, if sufficiently intense, mask an underwater sound signal. This noise should always be reduced below the level of the background arising in the water, when the latter is presented to the ear at a comfortable listening level. Since airborne noise is usually most intense at the low sonic frequencies, a greater signal intensity is required at low frequencies than at high frequencies.

Electrical noise may also mask a signal. With proper design and maintenance of the amplifying equipment, such noise can be made unimportant as long as a listening hydrophone is used whose minimum measurable pressure is below the pressure of the noise background present in the water. This is usually feasible, except possibly at high supersonic frequencies under quiet listening conditions.

The background of noise in the water may be reduced in importance by reliance on hydrophone directivity. The acoustically significant properties of such noise are as follows.

SPECTRUM

Peaks at certain frequencies are not frequent in background spectra and are generally unimportant when present. The sound power per cycle in the water falls off about 6 decibels each time the frequency is doubled, in much the same way as propeller cavitation, but increase of hydrophone directivity with frequency gives an effectively greater slope for the presented background spectrum than for the signal.

Modulation

The noise produced by the propellers of the listening ship shows marked modulation. Some noise shows marked peakiness, the peak factor increasing as increasing directivity or decreasing distance reduces the number of the individual noise sources contributing.

6.1.3 Signal-Background Mixture

The mixture of signal and background can be sharply modified by the receiving gear. The overall gain should be chosen to give a comfortable listening level (loudness level of about 70 decibels). Since signal recognition may occur at different frequencies for different targets, a wide-band system is most effective. Frequencies above 10 kilocycles may be heard most effectively if they are heterodyned down to sonic frequencies.

Some small advantage, especially in ease of concentration, is probably gained by use of a circuit which modifies the background spectrum to give equal loudness in a critical band at each frequency. In such a spectrum the sound power per cycle drops about 3 decibels per octave from 0.1 to 2 kilocycles, and is then constant up to about 6 kilocycles, rising at even higher frequencies. Limiting probably does not appreciably affect the signal recognition level but may make target identification more difficult.

6.2 RECOGNITION LEVELS

6.2.1 Steady Sounds

The masking of steady sounds is simplest when the noise background presented to the ear has a smooth spectrum, which falls off by not more than 20 decibels for each doubling of the frequency. Such a steady sound can be heard 50 per cent of the time at a particular frequency, when the signal level in one of the ear's critical bands centered at that frequency is just equal to the noise level in the same band. Thus a tone of frequency f_0 between 400 and 1,000 cycles can be heard when its sound power is equal to the noise power in a 50-cycle band centered at f_0.

The recognition level of a signal is that level at which the signal can be heard in at least one critical band. The recognition differential for the signal is then 0 decibel, if the signal and background are each measured in the critical band at which recognition occurs. If the signal and the background have parallel spectra, the *RD* is zero for any band used to measure the signal and noise. In other cases, the *RD* may be a large negative number if the signal and background are measured in a wide band. If the noise background has an irregular spectrum or falls off more rapidly than 20 decibels per octave, remote masking may occur, and sounds in one critical band may be masked by noise in another. This effect depends on the loudness level of the presented sound. Figure 6 in Chapter 2 should be used in this case.

6.2.2 Modulated Sounds

If a single-frequency peak fluctuates slowly (less rapidly than about 5 times per second) and if the noise has a gently sloping spectrum, the *RD* as measured in a critical band is still 0 decibel, provided that the peak level of the sound is measured. Modulation serves to call attention to signal components that might otherwise not be noticed; thus modulation produced by rotating a directional hydrophone helps to ensure most effective listening.

Modulated sounds whose spectra are parallel to that of the presented background and which are first heard in several critical bands simultaneously can be heard even when the peak intensity is below that of the background. The modulation may be inherent in the signal or may be imposed by rotation of a directional hydrophone. Provided that the modulation rate is between 1 and 10 per second, such a signal can be heard in half the trials if the total modulation of the signal-background mixture is 1 decibel. Thus if the peak signal intensity is used for computing the recognition level and measurements are made in a band over which the noise and signal spectra are parallel, the *RD* for a wide-band sound modulated 100 per cent is −6 decibels. For 50 per cent modulation the corresponding *RD* increases to −5.4 decibels, while for 30 per cent modulation it is −4.1 decibels.

6.2.3 Transition Curves

A plot of signal recognition probability against signal level is a transition curve. Such a curve shows not only the signal recognition level, defined for 50 per cent recognition, but also the levels required for other probabilities of recognition. The *spread* of such a curve is defined as the increase in signal level required to increase the recognition probability from 20 per cent to 80 per cent.

Variations in the spread from 1.3 to 8 decibels are observed. The spread tends to be less for cautious observers, but other changes in conditions have generally no systematic effect on the observed values of the spread.

6.3 TARGET IDENTIFICATION

Detailed identification of targets by their sound output may depend on the simultaneous perception of sounds at a number of different frequencies. Such perception may best be obtained by use of a wide-band listening system. Visual and mechanical methods are probably inferior to the ear in target identification, since they lack the filter properties provided by the ear's critical bands and since the quality of a sound cannot be simply recognized as yet by other than aural means.[a]

[a] N.B. Useful concepts, whose meaning is developed at diverse points in the foregoing and following chapters, have been indexed under unified headings to facilitate synopsis and review.

Chapter 7

CHARACTERISTICS OF ECHOES AND REVERBERATION

LISTENING IS SUBJECT to the serious limitation that a submarine can, if it wishes, be extremely quiet, and practically impossible to hear more than a few hundred yards away. Thus for reliable antisubmarine detection it is necessary to send out a pulse of sound and listen for the returning echo. This technique has the added advantage that it gives the distance, or range, of the submarine, information not readily obtained with underwater listening. In echo ranging as in listening, however, the signal can be masked by the background of sounds reaching the sonar operator. Knowledge of the audibility of echoes may be helpful in choosing the type of pulse to be sent out as well as in the design of the receiving gear. Hence studies of echo masking have operational importance, and considerable research along these lines has been carried out.

To understand the results obtained in this research on echo masking requires a knowledge of the basic properties of the signal and background, since from a basic standpoint these properties determine the recognizability of the signal. Some of these relevant properties have already been discussed in Chapter 3. For example, the properties of a noise background have been outlined in some detail in Section 3.2; since in many practical conditions noise may be the component of background which masks an echo, this previous discussion of noise is also relevant in echo masking studies. Similarly, the discussion of sonar gear given in Section 3.3 is also relevant here, since supersonic equipment used in echo ranging is frequently used also for supersonic listening.

The present chapter is therefore devoted to a description of several properties of underwater sounds which have not yet been described and which are important in the masking problem. First, in Section 7.1, some of the basic properties of short pulses are briefly discussed. The following section describes the quality of the echoes returned from various targets, and points out the amplitude and frequency properties of such echoes, which can be important in attempts to distinguish echoes from background. Finally, the properties of reverberation, which may be the dominant component of background if the noise level is low, are discussed in Section 7.3.

7.1 IDEAL PULSES

The basic sound used in echo ranging is a pulse of supersonic frequency. In practice, the upper limit on the length of the pulse is set by requirements of range accuracy and on the reverberation level, which increases with increasing pulse length. The lower limit is determined by deterioration of echo strength with decreasing pulse length and by the increasing difficulty of recognizing the shorter pulses. The pulse is generated electrically and is transformed into sound by the projector (transducer). The emitted sound wave then travels to the target and is reflected back. On its return to the hydrophone, or transducer, it is detected by the same equipment which generated it. Usually it is heterodyned down to an audible frequency and detected aurally. A variety of effects can modify the sound as it passes through the water, is reflected by the target, and is converted by the receiving equipment into an audible sound. However, the basic property of the original transmitted pulse is one of the important factors which determines the physical properties of the perceived echo. The character of such a supersonic pulse is therefore discussed in this chapter.

Ideally, a pulse consists of a train of waves of constant amplitude and frequency lasting for the time τ. While practical deviations from this ideal may exist and will be discussed later, analysis of this ideal pulse casts light on the physical situation.

7.1.1 Spectrum

While such an ideal pulse has a constant frequency during the limited time it is being generated, its effect on a tuned receiving system is not that of a sustained single-frequency tone. As a result of the short duration of the pulse, attempts to analyze the sound with sharply tuned systems show that the pulse is similar to a band of sound spread over a range of frequencies. This may be viewed on one hand as an effect of the transients produced in a tuned system by a short pulse. A system sharply tuned at any frequency will show some response for a very brief input of a quite different frequency. Thus a pulse only 1 cycle in length will produce response in any acoustical or electrical filter tuned to any frequency within a broad band; in this particular sense the pulse is indistinguishable from a noise pulse.

Another way of looking at the problem is to regard the short pulse as the sum of continuous sound waves whose amplitudes and phases are such that during the interval τ they add up to the pulse and at all other times they cancel out. This collection of sustained tones, equivalent to the pulse, is called the spectrum of the pulse; the amplitude of each tone of frequency f is denoted by $A(f)$. Each tone of amplitude $A(f)$ is sometimes denoted as a spectral component of the pulse. This analysis is convenient in that the amplitude of $A(f)$ gives immediately the response of a tuned filter of frequency f when the pulse is fed into it.

The amplitude of the spectral component may be found by methods which are standard in the theory of Fourier series and Fourier integrals. The sound pressure $p(t)$ in a short pulse of frequency f_0 is given by the expression

$$p(t) = p_0 \cos 2\pi f_0 t\,, \tag{1}$$

when the time t lies within the interval τ. If we measure the time from the middle of the pulse, then equation (1) holds when t is greater than $-\tau/2$ and less than $+\tau/2$. At all other times $p(t)$ vanishes. From this equation it may be shown that the amplitude of the Fourier component frequency f is given by

$$A(f) = p_0 \int_{-\tau/2}^{\tau/2} \cos 2\pi f(t) \cos 2\pi f_0 t dt\,. \tag{2}$$

This expression may be integrated directly and gives the relationship

$$A(f) = p_0 \left\{ \frac{\sin \pi (f + f_0)\tau}{2\pi (f + f_0)} + \frac{\sin \pi (f - f_0)\tau}{2\pi (f - f_0)} \right\}. \tag{3}$$

Several results are evident from this equation. In the first place, $A(f)$ is greatest when f is almost equal to f_0. Under these conditions the second term is very much greater than the first. The total energy in the spectral components present in a band 1 cycle wide is proportional to A^2. Most of this energy is at those frequencies for which the second term in equation (3) is large. As $f - f_0$ increases, the second term vanishes when the argument of the sine function is $\pm\pi$. Thus the value of f for which this term vanishes is given by

$$(f - f_0)\tau = \pm 1\,. \tag{4}$$

The total width Δf of this spectral region of high energy is obviously twice the value of $f - f_0$ found from equation (4). Thus

$$\Delta f = \frac{2}{\tau}\,. \tag{5}$$

While the spectrum of the pulse given by equation (3) extends to very large and very small values of the frequency, most of the energy of the pulse is included within the band whose width Δf is given by equation (5). This width is therefore called the *essential width* of the pulse spectrum. Alternatively, the "width" of the pulse spectrum is sometimes defined as the frequency separation of the two points where the spectrum level is 3 decibels down from its maximum value, between the points where $A^2(f)$ is $\frac{1}{2}A^2(f_0)$. On this definition, Δf is about equal to $1/\tau$. With this definition, the total energy in the pulse is just equal to the energy per cycle at the midfrequency f_0, multiplied by the essential width Δf, or $1/\tau$. In this report, however, equation (5) will be used to define the essential width Δf of the pulse spectrum.

If certain of the spectral components of the pulse are suppressed, the resultant sum of the different tones will no longer exactly resemble the pulse and will, moreover, contain some sound energy at time intervals outside the limits of the pulse. However, if all the frequencies in the essential width of the spectrum are included and those outside are omitted, most of the energy in the pulse will still be present although the shape of the pulse may be seriously distorted (see Figure 1 in Chapter 8).

This result, that the essential width of a pulse spectrum is inversely proportional to the pulse duration, lies at the base of much of the study of short pulses. Thus, a pulse which lasts only 10 milliseconds has an essential width of 200 cycles; a band-pass filter which admits less than 200 cycles or whose midfrequency is not centered at the pulse frequency f_0 will distort and weaken the pulse very seriously. Also when the response of tuned systems is used to determine the properties of such a short pulse, 100 cycles give an order-of-magnitude upper limit to the accuracy with which the pulse frequency can be simply measured. A receiving system admitting a band only 1 cycle wide would show about the same response to the pulse when tuned anywhere within a band some 25 cycles wide centered at the frequency f_0. To obtain very much greater accuracy with tuned systems, it would be necessary to measure $A(f)$ for a number of different values of f and to determine f_0 as the midfrequency of the observed spectrum.

Another important result also follows from equation (3). This equation may be used to give the energy per cycle at frequencies outside the essential width of the spectrum. The energy per cycle is proportional to A^2; outside the essential width of the spectrum the sine functions of equation (3) will oscillate between $+1$ and -1, but on the average, the square of the sine equals $\frac{1}{2}$, while the average value of the product of the two sines is zero. Thus if we average over bands which are wide compared with the essential width of the spectrum, but centered far away from the midfrequency f_0 of the spectrum, we obtain

$$A^2(f) = \frac{p_0^2}{4\pi^2} \frac{f^2 + f_0^2}{(f^2 - f_0^2)^2}\,. \tag{6}$$

From this equation it follows that when f is not too widely different from f_0, the numerator in equation (6) is constant, and the average energy per cycle decreases inversely as $(f-f_0)^2$. For a frequency f much greater than f_0, the average energy per cycle is proportional to $1/f^2$. In this situation the spectrum level decreases 6 decibels every time the frequency is doubled. It will be noted that this level is independent of the pulse length. Thus, even for a long pulse, energy will be present at frequencies far removed from the midfrequency f_0. However, the total energy in the pulse increases directly with the pulse length, and thus the relative amount of the energy at such distant frequencies decreases with increasing pulse length as expected.

This presence of frequency components at frequencies far removed from the midfrequency f_0 depends very critically on the abrupt beginning and termination of the assumed ideal pulse. When the pulse begins more gradually, building up to full amplitude in several cycles and falling off similarly at the end, the essential width of the spectrum is not appreciably changed, but equation (3) can no longer be used to find the energy at frequencies far removed from f_0. In fact, the energy present at such frequencies will be very much less for a rounded pulse than for an ideal square-topped pulse.

The importance of these results, summarized in equations (5) and (6), is emphasized by the fact that the distribution of energy in the spectrum of a pulse is not changed by most of the processes which modify the pulse as it travels out and is reflected, received, and finally recognized as an echo. In the first place, most of the phenomena affecting a pulse are linear; this means that their magnitude is independent of the absolute intensity of a pulse, and an emitted pulse with twice the original amplitude will be exactly the same as the original pulse when it is finally presented to the operator, except that its amplitude will be twice as great. As a result of this fact, each spectral component is unaffected by the presence of others and is transmitted and reflected in exactly the same way as would a single continuous tone of that frequency. In the second place, both the transmission and the reflection of sound energy do not change much with frequency over the essential spectrum of a supersonic pulse, provided that this is not much less than a millisecond in duration. As a result of these two conditions, the energy in each 1-cycle band in the spectrum of the echo returning to the transducer should be the same as the energy in each 1-cycle band in the spectrum of the transmitted pulse. The phases of different spectral components will be modified as the pulse is transmitted and reflected, thus leading to distortion of the pulse, but the magnitude $A^2(f)$ should be unaffected by these processes. Thus in the returning echo, as in the pulse, the width of the essential spectrum is given by equation (5), and for high frequencies the energy per cycle should decrease 6 decibels for each doubling of the frequency. The receiving equipment if sharply tuned may modify the distribution of energy among the different spectral components and thus change the spectrum of the echo presented to the sound operator. If the hydrophones or loudspeaker used are not flat over the pulse spectrum, the distribution of energy over the spectrum may again be changed. These effects are usually important only when very short pulses are considered (less than 10 milliseconds long), or when spectral components far from the midfrequency are important.

The foregoing considerations have been developed for a pulse with certain ideal properties. Actual echo-ranging projectors do not usually generate such ideal pulses. A projector which is mechanically resonant, such as a magnetostriction projector, is very inefficient for sustained tones whose frequencies are far from resonant. For this reason, if an ideal square-topped pulse is fed electrically into such a projector, the distant-frequency components will not be radiated and the pulse will therefore become rounded, since these distant components are necessary to produce the abrupt beginning and ending of the pulse. This effect is identical with the distortion of the spectrum of the received echo by a tuned receiving system discussed earlier in this section. Other imperfections in practical equipment may be expected to produce some irregularities both in the amplitude and the frequency of the pulse, and thus to modify its spectrum from the simple form given in equation (3). All these possibilities must be kept in mind when the observed spectrum of a pulse is compared with theoretical expectations.

7.1.2 Frequency Modulation

The frequency properties of emitted pulses have been studied by means of the periodmeter,[1] a device which measures the time interval between successive zeros in the sound pressure. This interval may be regarded as the period of oscillation of the wave. Thus, the instrument can be used to determine very accurately the instantaneous frequency of a sound wave, defined as the reciprocal of the measured period. If the successive periods are not equal, the periodmeter then gives information on the distribution of these periods, that is, on the number of intervals lying within each range of time. This information can be interpreted as giving the distribution of frequencies in the emitted pulse, since to each period there corresponds an instantaneous frequency. This information is quite different from that provided by an analysis of the pulse spectrum. For example, in an ideal pulse the emitted sound pressure is an exact cosine function, which equals zero at regular intervals and which, therefore, has a constant period.

The periodmeter would show that such a pulse had a constant frequency for its entire duration. The spectrum, however, would show a spread of frequencies in accordance with equation (3). While this difference between the results of the periodmeter and the spectrum must be kept in mind, the periodmeter provides use-

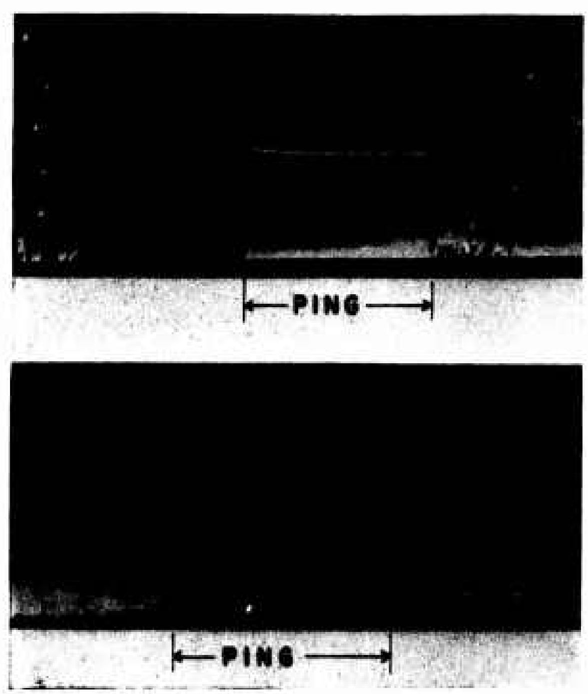


FIGURE 1. Periodmeter analyses of pulses. Horizontal lines indicate 20-cycle intervals.

ful information on the basic frequency properties of sound waves.

When the periodmeter is used to analyze pulses sent out by echo-ranging gear, variation in the frequency of the emitted sound is often discovered. Typical examples are shown in Figure 1, where periodmeter records are shown for pulses produced by special laboratory electronic gear at University of California Division of War Research [UCDWR] (lower record) and by standard Navy sonar gear on board one of the West Coast Sound School ships at San Diego (upper record). The height of each vertical white line represents the interval between two successive zeros in the sound pressure. Since the scale is nonlinear, horizontal guide lines have been drawn at intervals corresponding to 20 cycles in the instantaneous frequency. The record to the right of each ping represents reverberation but is garbled by chattering in the change-over relay. The durations of the pulses shown were about 50 milliseconds. These records were made from the signal input to the transducer, but similar results were obtained from the received acoustic signal.

The pulse obtained with the high quality equipment shows no frequency modulation, but the pulse produced by the standard gear shows a variation of frequency over a range exceeding 40 cycles. Although this variation is not large, it could be important in some situations. Especially for longer pulses, for which the essential width of the spectrum is narrow, the presence of frequency modulation in the actual pulse can significantly broaden the spectrum of both the pulse and the echo (see Figure 2). This effect is particularly important in the accurate determination of doppler shifts.

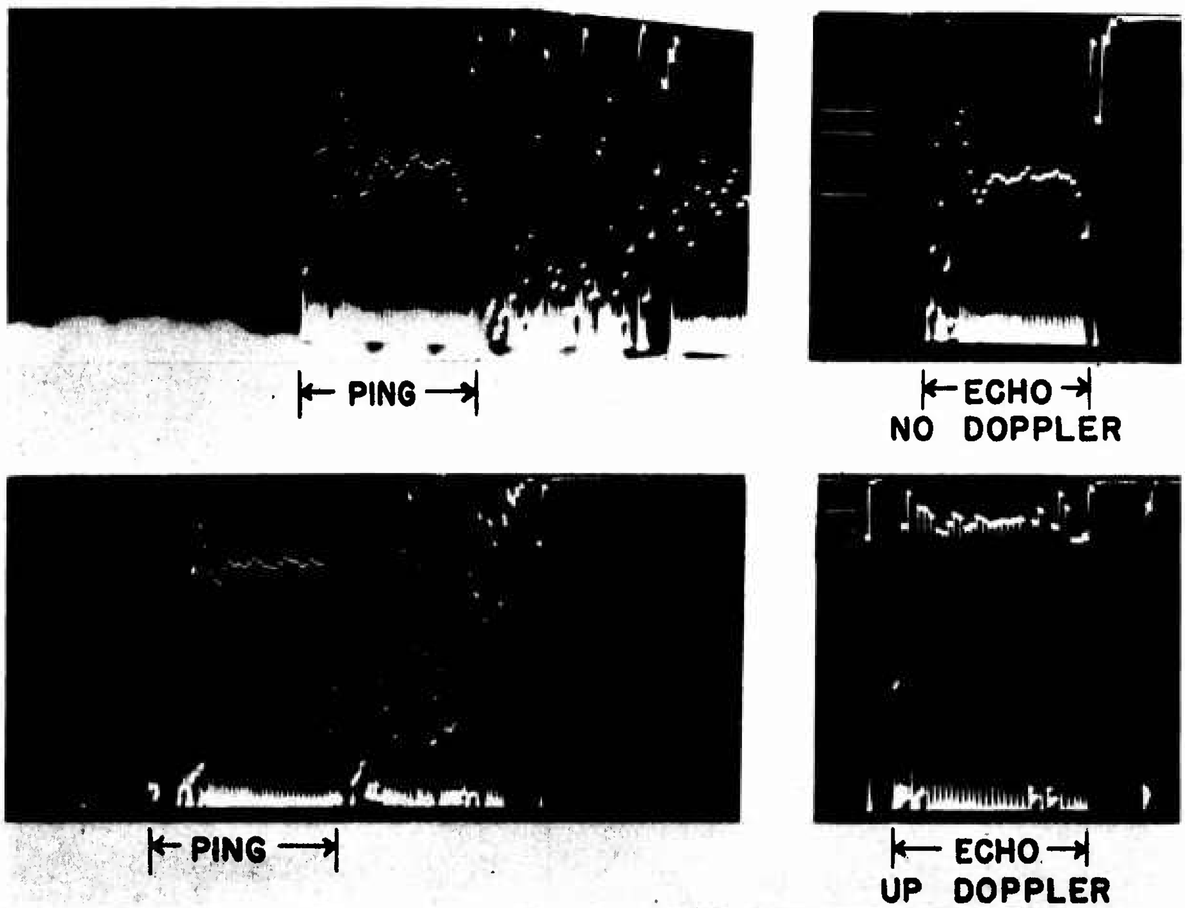


FIGURE 2. Periodmeter analyses of pulses and echoes. Horizontal lines indicate 20-cycle intervals.

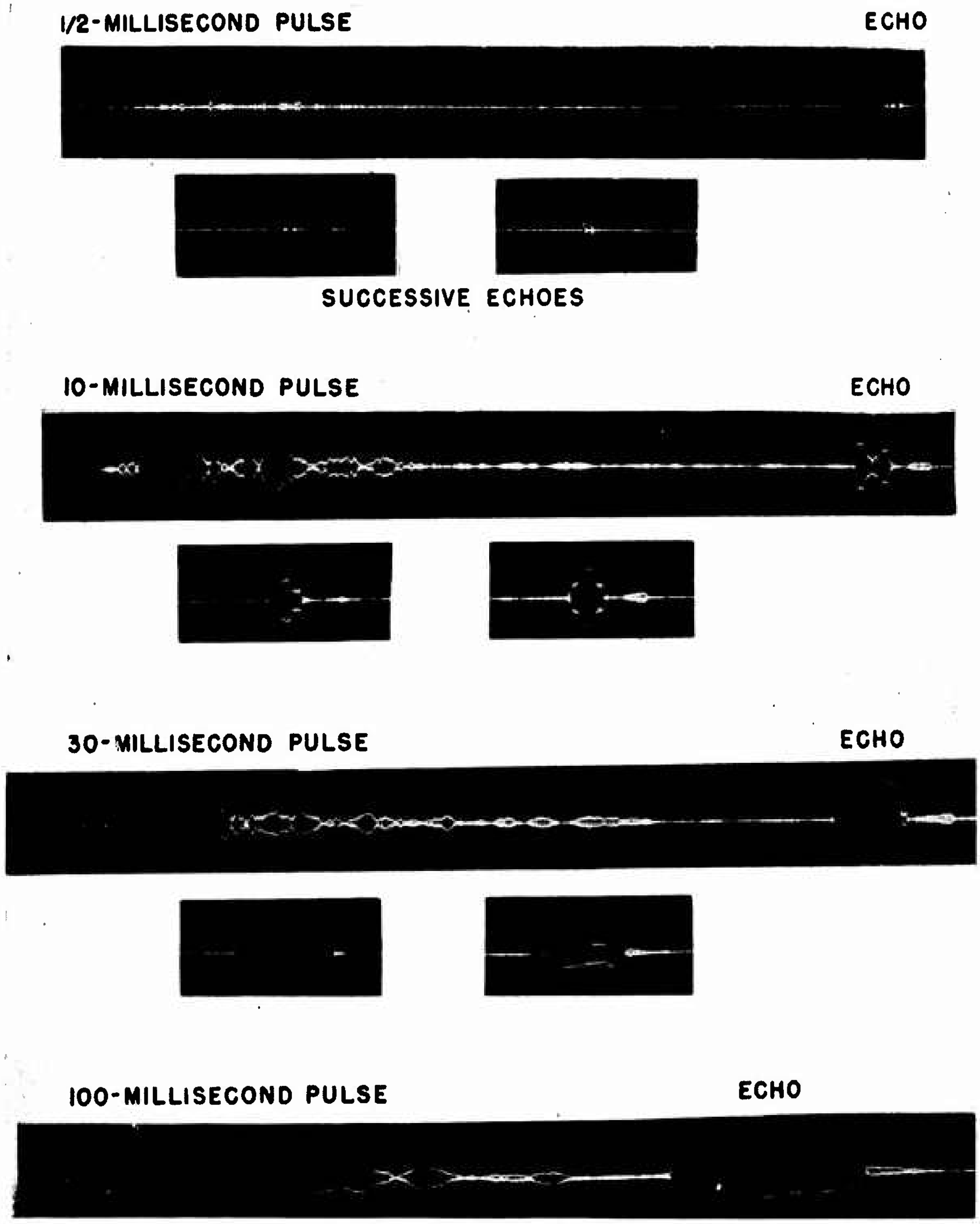


FIGURE 3. Echoes from a submarine at beam aspect.

7.2 ECHOES

An echo can, in general, be much more complicated than the pulse which produces it. The sound may become distorted during transmission through the water, especially in shallow water where multiple paths tend to prolong the signal received at the target. In addition, the sound is distorted by reflection from the target. This second effect is usually more important than the first, and will be given chief attention here. The amplitude and frequency characteristics of observed echoes are discussed in the following two sections.

7.2.1 Amplitude

A small geometrically smooth target, such as a sphere or triplane, reflects the sound specularly back toward the echo-ranging projector. Such specularly reflected sound is generally a relatively exact reproduction of the received pulse. In practice, the echo from a mine case may be of this character. Echoes from large targets, however, such as surface vessels and submarines, are generally the sum of echoes from many reflectors. While the echo from each reflecting surface may reproduce the pulse, the sum of the component echoes sometimes bears relatively little resemblance to the pulse. For example, the phases of different component echoes will interfere constructively with each other, while in other cases, the interference will be destructive. Thus, in general, the total received echo may be considerably prolonged from the original short pulse and may contain considerably more amplitude modulation, resulting from interference between the different component echoes making up the total echo.

The simplest echoes from large targets are those received from a ship at beam aspect. For a submerged submarine at beam aspect, for example, the echo comes primarily from specular reflection produced by the cylindrical sides of the submarine, from the pressure hull and ballast tanks. This echo, like that from a sphere, tends to be a fairly accurate reproduction of the pulse, as is evident from the traces shown in Figure 3. It is evident from the reproduction for a short pulse shown in this figure that even these beam echoes show some structure. Thus several reflecting surfaces produce the observed beam echo.[a] For surface vessels, the available evidence is not complete but indicates that the echo returned at beam aspect tends to be a fairly accurate reproduction of the outgoing pulse. Such an echo, which like the pulse is approximately square-topped, is sometimes known as a *clean* echo, in contrast to the *smear* echo discussed in the next paragraph.

Observations on submerged submarines show that specular reflection from the sides of the submarine begins to weaken when the axis of the submarine is not exactly perpendicular to the sound ray. When the submarine aspect differs by more than 15 degrees from the exact beam, this specularly reflected sound presumably goes off in another direction and is not detected back at the echo ranging vessel. The echo which is observed at these off-beam aspects is apparently produced by reflecting and scattering surfaces distributed over the entire length of the submarine. Thus, the echo will be considerably prolonged over the pulse duration τ by an amount depending on the length of the submarine in the line of sight. The observed duration of the echo, in seconds, is equal numerically to $\tau + (L \cos \theta)/2c$, where L is the length of the submarine, θ is the angle between the submarine axis and the direction of the sound ray, and c is the velocity of sound.[a] Such an echo, which is prolonged relative to the pulse and whose envelope is jagged and uneven, has been called a *smear* echo in contrast to the *clean* echo discussed previously. The average intensity in a smear echo would be expected to decrease with decreasing τ since the total energy of the pulse, and therefore of the echo, must decrease as τ decreases. It has been observed experimentally that the peak intensity of a smear echo tends to decrease with τ especially when τ is less than 10 milliseconds.[a]

Sample smear echoes for different pulse lengths are shown in Figure 4. All these were obtained from a submarine at off-beam aspects. It will be noted that the structure of these off-beam echoes begins to resemble that of reverberation, discussed in the following section. In particular, the length of each blob observed in the echo is roughly the same as the length of the emitted pulse. It is uncertain whether the heights of the different blobs observed in these smear echoes are correlated with different re-

[a] For a more detailed discussion see Division 6, Volume 8, Chapters 23 and 24.

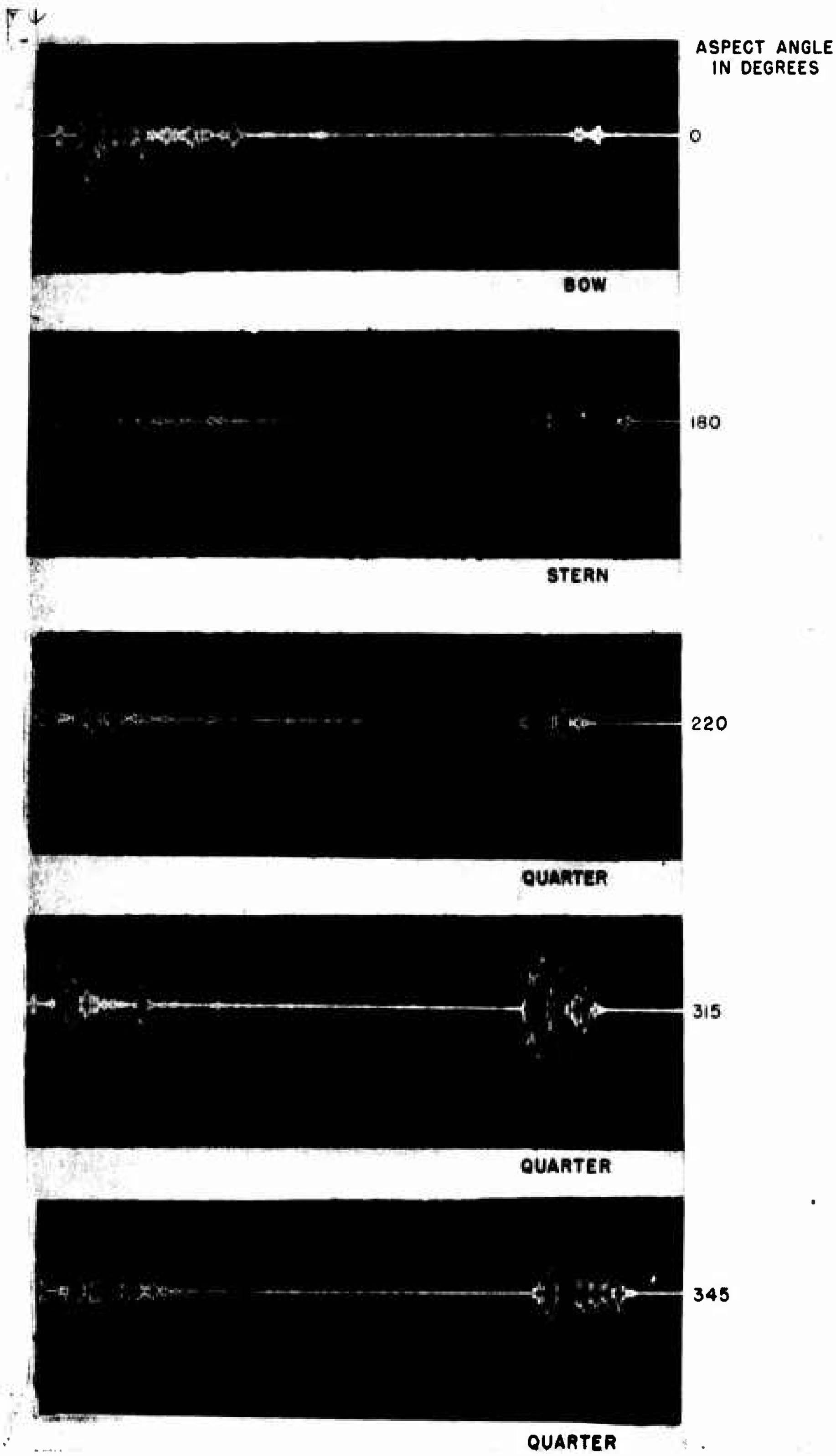


FIGURE 4. Echoes from a submarine at off-beam aspects.

flecting surfaces on the submarine or whether they are entirely the result of random interference between all the different scatterers. The extent to which the structure of a smear echo tends to repeat itself from one echo to the next is therefore uncertain. However, there is no question but that the time-amplitude pattern of such a smear echo tends to be irregular, with the blob size roughly equal in length to the pulse duration. While detailed investigations have been carried out only for submarines, these same conclusions are probably valid for any type of large target.

7.2.2 Frequency

The frequency properties of the returned echo are partly determined by the frequency properties of the pulse. In particular, it has already been noted that the energy per cycle in the spectrum of the returning echo is almost exactly the same as that in the spectrum of the emitted pulse. If the target is in relative motion, the entire spectrum of the echo will be shifted by the frequency $(2v/c)f_0$, where v is the motion of the target in the direction of the incident and reflected sound, c is again the velocity of sound, and f_0 is the midfrequency of the pulse. This doppler shift simply translates the entire spectrum by a uniform amount, however, and does not otherwise change the signal. Motion of the echo-ranging projector through the water also produces a doppler shift in both the emitted signal and the received echo. These shifts are not primarily important in masking studies, however, since the doppler shift of interest is that of the echo relative to the reverberation, which is given by the preceding formula.

While the distribution of energy in the echo spectrum is identical with that of the pulse, apart from a uniform doppler shift, the relative phases of the different spectral components may be considerably modified by interference between the different component echoes making up an observed smear echo. These interference effects in a smear echo can affect the instantaneous period of the sound. Results obtained by the periodmeter show that smear echoes can possess a considerable amount of frequency modulation. Since the factors governing the response of the ear to complex and variable sounds are not yet well understood, it is uncertain whether the distribution of instantaneous frequencies shown by the periodmeter or the different frequency components shown by the spectra correspond most closely to the properties perceived by the ear. It is possible that both of these methods of describing echoes must be taken into account in a realistic explanation of how the ear recognizes echoes.

Echoes can also be obtained with pulses of frequency-modulated sound. One situation of practical importance is that in which the emitted pulse has a constantly ascending frequency. As with a pulse of constant-frequency (CW pulse), a frequency-modulated (FM) pulse striking a vessel at beam aspect gives an echo which tends to reproduce the outgoing pulse. For example, echoes from a submarine at beam aspect, obtained with the pulse shown in Figure 2, show exactly the same frequency modulation as that in the original pulse. Off-beam echoes, on the other hand, will again involve the superposition of the different component echoes from the different reflecting surfaces. With FM pulses, this combination of component echoes from different ranges will combine sounds of different instantaneous frequencies, and the resultant echo may be expected to have rather complicated characteristics. These effects have not been studied in detail, but may be expected to affect the performance of the ear in recognizing such echoes. In addition, the intensity of echoes produced with an FM pulse seems to show somewhat less variability from echo to echo than is found with CW echoes.

7.3 REVERBERATION

If the ocean were perfectly uniform and the surface and bottom were smooth and flat, the only sound reflected back to the echo-ranging projector would be that from the target, and the noise level would be the only source of masking. In fact, however, the ocean is not uniform and the ocean surface and bottom are not perfectly smooth and flat. As a result, sound is scattered from the many inhomogeneities in the body of the ocean and from the irregularities at the surface and bottom. This scattered sound is heard back at the echo-ranging installation as a rolling sound known as reverberation. At close range the reverberation level is high and will usually lie so far above noise level that it forms the dominant part of the background against which an echo at short range must be recognized. Even at longer ranges reverberation under some conditions, as for instance, over a shallow rocky bottom can be the important component of the background. The reverberation intensity is proportional to the total sound power radiated into the water, and thus increases proportionally with the pulse length, as long as the pulse length, in yards, is much less than the range at which the reverberation is measured.

Since reverberation is the sum of the individual echoes returned by myriads of scatterers, it is very similar in its properties to the smear echoes discussed in the preceding section.[b] Those features which are important to the masking problem are summarized here. First, the characteristics of reverberation produced by a single-frequency, or continuous-wave pulse (CW reverberation) are discussed. Secondly, the available evidence on the reverberation from a frequency-modulated pulse (FM reverberation) is briefly summarized.

7.3.1 CW Reverberation

The reverberation received from a CW pulse has approximately the same frequency as the outgoing pulse. However, its amplitude is modulated in an irregular way, and a closer examination of its frequency pattern shows irregularities there also. The amplitude properties of reverberation are discussed in the next section, while the frequency properties are treated later.

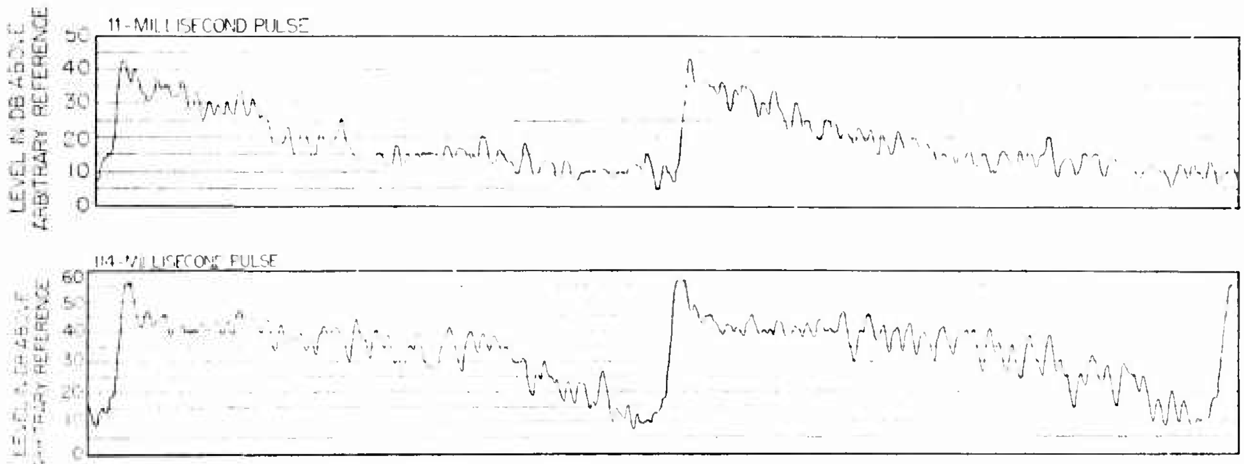


FIGURE 5. Time-amplitude-pattern of CW reverberation showing the variability in reverberation received from successive transmissions under nearly identical conditions. The writing speed was higher for the shorter pulse. For a better indication of the relative blob durations see Figure 7 in Chapter 9.

Amplitude Modulation

The amplitude variation of CW reverberation is very marked. Typical power level records of reverberation generated by different pulse lengths are reproduced in Figure 5. These were made from an oscillograph responding directly to the received reverberation at 24 kilocycles before heterodyning. It was shown by Lord Rayleigh many years ago that, when sound is received from many scatterers, the probability $P(I)$ of receiving an intensity greater than I is given by

$$P(I) = e^{-I/I_0}, \tag{7}$$

where I_0 is the average intensity received.

[b] The detailed properties of reverberation are discussed in Division 6, Volume 8, Chapters 11 through 17.

This distribution of intensity is known as the Rayleigh distribution. Equation (7) may be derived from the consideration of the random phases of the scattered sound.[c] In the case of reverberation, the intensity I_0 is itself a function of time and must be found by measurements over a great many successive reverberations. In each reverberation the intensity must be measured at a fixed time interval t after the pulse has been emitted. The average of all these intensities over many successive reverberations is then called the average intensity at the time t. The probability of deviations from this average value is given by equation (7). This equation has been confirmed experimentally for CW reverberation and may be assumed to be correct for all frequencies of practical interest. While these studies were carried out for the unheterodyned reverberation, the same result would presumably apply to the heterodyned reverberation also.

The blob size in the unheterodyned reverberation shown in Figure 5 is roughly equal to the length of the emitted pulse, a result which is also in agreement with theoretical expectations.

Decay Rate

The rate at which reverberation decays can have an important influence on the problem of echo recognition. It is customary to plot a decay curve for reverberation showing how the average reverberation I_0 changes with time elapsed since the emission of the pulse. In inspecting such a curve, it must be borne in mind that deviations from the mean intensity have the probability given in equation (7) and that actual reverberation is consequently much more jagged and irregular than is shown by the mean curve. The rate at which the average reverberation decays is found to be a highly variable quantity. Sometimes the reverberation falls away rapidly to a very low level; sometimes it remains at a sustained level for several seconds and then gradually diminishes. In shallow water, for example, when temperature gradients bend a sound beam downward, the reflected sound received from the bottom appears quite suddenly and for a relatively short time produces a crash of reverberation which may, under some conditions, be mistaken for an echo. This variability of the decay rate of reverberation, as oceanographic factors vary, makes it difficult to devise automatic level stabilizers which will keep the background at a constant level without smoothing out the echo.

Frequency Properties

One of the most noticeable characteristics of CW reverberation is its tonality. Since reverberation is simply the sum of many echoes, the distribution of the energy in its spectrum is the same as that in the spectrum of the emitted pulse. The scattering centers are usually not moving through the water at any high speed; therefore, reverberation usually has no doppler except that produced by the motion of the echo-ranging projector or receiver. The only exception to the latter statement is encountered in the case of strong currents, which produce systematic shifts between the reverberation from the water and the reverberation from the bottom. Because water reverberation is likely to come in at close range and bottom reverberation at longer range, a marked and systematic shift of reverberation frequency with range may be produced. Such effects will, of course, shift the entire spectrum without modifying the relative amplitudes of the spectral components.

Motion of the echo-ranging projector can produce marked doppler shifts in the received reverberation. Since these frequency shifts vary with relative bearing, and since the projector has an appreciable beam width, the spectrum of reverberation may be appreciably broadened by the motion of the echo ranging vessel. A detailed discussion of this effect is postponed to Section 10.12, and in the present description only the reverberation received by a stationary transducer is considered.

An analysis of the spectrum of the reverberation obtained with a stationary projector was made by the British early in the war. By use of a narrow-band filter, the reverberation energy per cycle at various frequencies was

[c] See Division 6, Volume 8, Chapter 7.

determined. This energy distribution was found to be identical with that measured for the pulse, in agreement with expectation. However, a similar investigation at UCDWR, making use of the periodmeter, yielded results not in agreement with theory.[1] Further observations are therefore required before the theory can be regarded as confirmed.

It will be recalled that the periodmeter gives information on the instantaneous period of the reverberation, which is not directly related to the spectrum. A typical result is shown in Figure 6, which is a reproduction of a periodmeter record for reverberation. As before, the height of each vertical white line in each of the diagrams represents the time interval between successive zeros in the heterodyned reverberation; the corresponding frequency scale is indicated by the horizontal white lines.

The pulse producing these reverberations showed very little frequency modulation, since laboratory equipment was used to produce a pulse similar to that shown in the upper record of Figure 1. Nevertheless it is evident from Figure 6 that the distribution of zeros in each reverberation sample is highly irregular, with rapid fluctuations from one period to another. An analysis of other records shows that for reverberation from a 92-millisecond pulse the root mean square deviation of periods is about 2 per cent of the average period, corresponding to a frequency deviation of 16 cycles at 800 cycles. While the reverberation was obtained from a projector moving at 8.5 knots, it was found that this ship speed was too low to produce any marked effect on the observed spread. The corresponding spread found for the outgoing pulse is, of course, very much less, and for an ideal pulse, with no frequency modulation, would be zero.

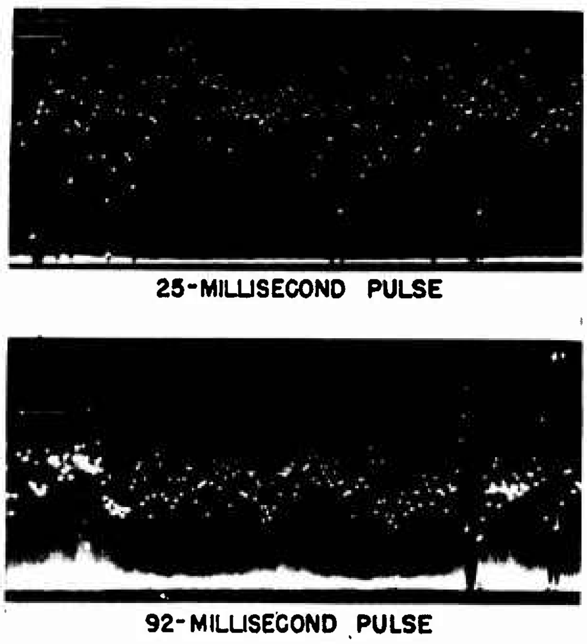


FIGURE 6. Periodmeter analyses of volume reverberation. Horizontal lines indicate 20-cycle intervals.

This spread of instantaneous periods in the received CW reverberation is the direct result of the severe amplitude modulation of the reverberation. While the detailed nature of the periodmeter results has not been explained, it seems clear that the interference of echoes from many targets produces a sound which in effect has rapid frequency modulation as well as amplitude modulation. This spread of instantaneous frequency, as well as the width of the essential spectrum, probably affects the performance of the ear when presented with these sounds.

Systematic changes of reverberation frequency with time have also been observed with the periodmeter. The observations indicate that the reverberation which arrives immediately after the pulse is sent out may have a higher or lower frequency than the reverberation arriving several seconds later. It has already been pointed out that some such effect would be expected in the presence of variable water currents, especially in shallow water. However, a steadily rising or falling pitch in reverberation seems to occur more frequently than can be explained on this basis. Possibly these changes in pitch are simply the result of statistical fluctuations. In any event, systematic frequency changes during the arrival of reverberation provide one more factor that may influence the masking effect of reverberation on echoes.

Heterodyne Frequency

For aural recognition, the reverberation and echo are heterodyned down to an audio frequency, which has generally been in the neighborhood of 800 to 1,000 cycles. It is therefore relevant to inquire as to the possible effects which heterodyning may have on the background. The spectrum will, of course, be largely unaffected, and the energy per cycle will have the same distribution about the new sonic midfrequency f_0' as it originally possessed about the supersonic midfrequency f_0. Similarly, the amplitude characteristics of reverberation would not be expected to be influenced appreciably by the heterodyning process.

FIGURE 7. CRO trace for reverberation from 271-millisecond pulse heterodyned to 800 cycles.

A CRO trace of heterodyned reverberation is reproduced in Figure 7. Comparison of such traces with those obtained for unheterodyned reverberation recorded at the same time shows that the abrupt reversals of phase found at various points in Figure 7 are much more common in the heterodyned sounds. Theory indicates that this effect is especially likely when the amplitude is relatively low. Such phase changes may conceivably have an appreciable effect on the masking properties of the reverberation.

7.3.2 FM Reverberation

When a pulse of rising or falling frequency is sent out into the water, the returning reverberation loses the marked tonality which it possesses for CW pulses. Echoes from scatterers at different ranges have different frequencies at any one time, and the combination of these different echoes produces a sound which is similar to a wide-band noise. The spectrum of the reverberation will be similar to that of the pulse, although over a region many thousands of cycles wide, exact similarity is not to be expected. While the total intensity of FM reverberation is the same as that of CW reverberation, the energy per cycle will be less, since the same energy is spread over a much wider band of frequencies. The energy of the echo per cycle is similarly reduced, of course.

Reverberation from an FM signal is also characterized by a different type of amplitude modulation than is found for CW reverberation. In the first place, the variability from the mean intensity I_0 at any time, as determined by measurements in successive reverberations, is less. For the Rayleigh distribution, applicable to CW reverberation, the rms deviation of reverberation amplitudes is 52 per cent of the mean amplitude. For FM reverberation, on the other hand, the observed rms deviation is only 33 per cent of the mean reverberation amplitude.

In addition, the rate of variation is much greater for FM reverberation than it is for CW reverberation. In the latter case, the typical reverberation blob has about the same duration as the pulse, provided that the pulse duration is much less than the elapsed time between the emission of the pulse and the arrival of the reverberation which is to be measured. Since the duration of the echo tends to be about the same as that of the pulse, except for very short pulses and extended targets, a blob of CW reverberation tends to have about the same shape as an echo. This naturally increases the difficulty of echo recognition. With an FM pulse the duration of the echo is unchanged but the length of a typical reverberation blob becomes very much less. This difference between FM reverberation and echo makes it possible to smooth out the reverberation without also smoothing out the echo; the echo thus stands out clearly above the smoothed reverberation. Similarly, the aural detection of an echo without doppler presented against a reverberation background, is improved somewhat by the use of FM. The precise extent of this improvement is uncertain, since no detailed measurements have been made. The loss of doppler discrimination with FM pulses is such a serious disadvantage, however, that FM signals are primarily useful in such applications as mine detection. In this case, the small size of the target makes it possible to use very short signals to reduce reverberation, without reducing the strength of the echo. With such short pulses and with no doppler present, visual detection is probably preferable to aural in any case, and full advantage may be taken of FM pulses.

Chapter 8

NOISE MASKING OF ECHOES

Most of the data described in this chapter relates to the recognition of artificial and of recorded echoes, from CW pulses, in the presence of noise (usually, thermal noise). As in Chapter 4, observations are grouped initially according to place of origin, because of differences of procedure followed at various laboratories, and summaries of general trends are given later in the discussion.

8.1 BELL TELEPHONE LABORATORIES TESTS

Studies have been made of a number of factors which may affect the practical situation, such as pulse duration, heterodyne frequency, system band width, and listening level, as well as various modulations, distortions, and time patterns. Most of the general ideas developed during the course of the previous analysis apply also to the masking of pulses. In Chapter 2, for example, the width and, indeed, the concept of a critical band have been defined for relatively sustained tones only. These definitions seem to require no substantial modification even for fairly short pulses; in other words, differences in pitch of echo and background components provide a helpful cue. On the other hand, recognition of tonal pulses would be expected to deteriorate with diminishing pulse length since the loudness of pulses is a function of their duration (see Figures 12, 13, and 14 in Chapter 2), and loudness, or degree of neural stimulation, is related to masking. In addition, chance noise peaks which coincide with, or slightly precede the echo cut its effective duration; hence, the detectability of pulses should continue to improve even when their lengths exceed 250 milliseconds (see Figure 8). For this reason, the "squareness" of the echo envelope (the rates of growth and decay at the leading and trailing edges) should help the listener, and any process (such as constrictive filtering or envelope distortion produced by phase interference during transmission of the underwater sound) which diminishes differences between the time-amplitude patterns of the echo and the noise peaks should hinder detection. Conversely, distinctive modulations of the echo may be expected to aid detection. These, and similar points are discussed here in connection with the experimental evidence.

8.1.1 General Procedure

In these tests,[1] the masking background was a band of thermal noise extending from 200 to 3,000 cycles and essentially flat over that interval, which is a fair approximation to the practical situation (see Figures 66 through 75 in Chapter 4). Fluctuations from the mean level of the noise in the 2,800-cycle band, as measured with a thermocouple-ammeter arrangement, were occasionally as great as ± 1.5 decibels, but on the average did not exceed ± 0.5 decibel. The test apparatus was essentially that shown in Figure 1 in Chapter 4, except that a pulse generator was used to feed the signal channel. The response of the headphones was very nearly flat within the limits of the 2,800-cycle noise band. Tests were administered to groups of between 5 and 12 observers seated in a quiet room. Gain in the background channel had the same value in all the tests, but the gain applied to the output of the mixture amplifier could be adjusted by the individual observers to give the level each considered optimal for performance and comfort.

Each test consisted of 20 signals, and during the time required to complete any test the noise background was presented without interruption, that is, without subdivision into an alternating series of listening periods and quiet intervals. The signal strength and thus the signal-to-background ratio were held constant for each group of 20 signals, which were injected according to a random time pattern in which the average interval between successive signals was about 8.5 seconds, although the spacing between two signals was occasionally as little as

2 seconds or as much as 15 seconds. The observers were instructed to press a telegraph key momentarily, as soon as they thought they heard a signal, in order to encourage a responsive rather than a cautious attitude. The voting key actuated a counter designed to register separately all errors of omission and commission. Failure to depress the key within 1 second after the injection of a signal was automatically scored as an omission, while a later vote (made at any time before the next signal) was scored as an error of commission; thus, random guesses were about 7.5 times more likely to be registered as commissive errors. The arrangement of the counting circuit precluded the possibility that a long vote would register in more than one category. The fact that the results obtained in these tests are in good agreement with data obtained in tests involving a quite different scoring technique (see Sections 8.2 and 8.4) shows that a listener can decide and indicate within 1 second whether or not a signal is audible, at least, when the probability of perception is 50 per cent or better and when no distinction need be made between one kind of signal and another (such as wake echoes and submarine echoes).

A series of tests was made for each type of pulse, the highest signal-to-noise ratio being used first in the series; a rest period was given after each test to minimize fatigue. The signal-to-noise ratios used in successive tests were diminished in steps of 2 to 3 decibels so that the range between approximately 100 per cent and approximately 50 per cent recognition probability was covered in a series of 4 tests. In order to give some weight to errors of commission, recognition probability was defined as $(N-E_{-})$ $(N+E_{+})$, where N is the number of signals presented, and E_{-} and E_{+} indicate the number of errors of omission and commission, respectively, so that occurrence of a substantial number of commissive errors would significantly reduce the calculated value of recognition probability. Since errors of commission were relatively infrequent (see Figure 2), values computed from the preceding expression were very nearly equal to those which would be obtained by the procedure described in Section 8.2, that is, by rejecting all tests in which E_{+} is significantly greater than zero and defining recognition probability by means of the fraction $(N-E_{-})/N$. The transition curves obtained for the various kinds of signal and background used in these tests are discussed in the following section.

8.1.2 Effects of Pulse Length and Band Width

The intrinsic frequency of the CW pulses which were studied was about 740 cycles; their durations, 600, 200, and 67 milliseconds. Three filters (with nominal band widths of 15, 5, and $1\frac{2}{3}$ cycles, respectively) were used to restrict the frequency range of the interfering background. It will be observed that the band widths selected are very nearly the smallest required (twice the reciprocal of the pulse length in seconds; see equation (5) in Chapter 7) to transmit rectangular pulses of the indicated durations without seriously reducing their energy. Since, in practice, pulses and interfering waterborne backgrounds are necessarily transmitted to the ear over the same circuit, it is important to pass signal and noise through the same filter in order to avoid introduction of spurious effects in this type of test; see Section 9.1.2. This procedure was followed where necessary. The output envelopes of the originally nearly rectangular pulses were examined with the aid of a persistent cathode-ray oscilloscope [CRO] screen, and the observed shapes are reproduced in Figure 1. As indicated in Section 4.2.3, noise peaks are prolonged by the action of constrictive filters so that the minimum duration of each peak approaches twice the reciprocal of the filter width.

The transition curves in Figure 1 give the observed recognition probabilities for several combinations of pulse length and band width. The abscissas refer to the ratios of the overall levels measured in the signal and noise channels separately. Since all results are referred to a common basis (level of the 2,800-cycle noise band), the amount of improvement produced by a given change in the test conditions may be obtained by inspection. Since no observations were made at signal-to-noise ratios giving substantially less than 50 per cent recognition probability, the abscissas have been drawn at the latter value for the sake of convenience in reading the curves.

Several aspects of the data shown in Figure 1 are worth examining. To begin with, it may be asked whether the width of a critical band, for the case of a tonal pulse masked by wide-band noise, is equal to the width measured for a sustained tone of the same intrinsic frequency as the pulse. Use of the direct approach, of the type discussed in Section 2.3, would imply that

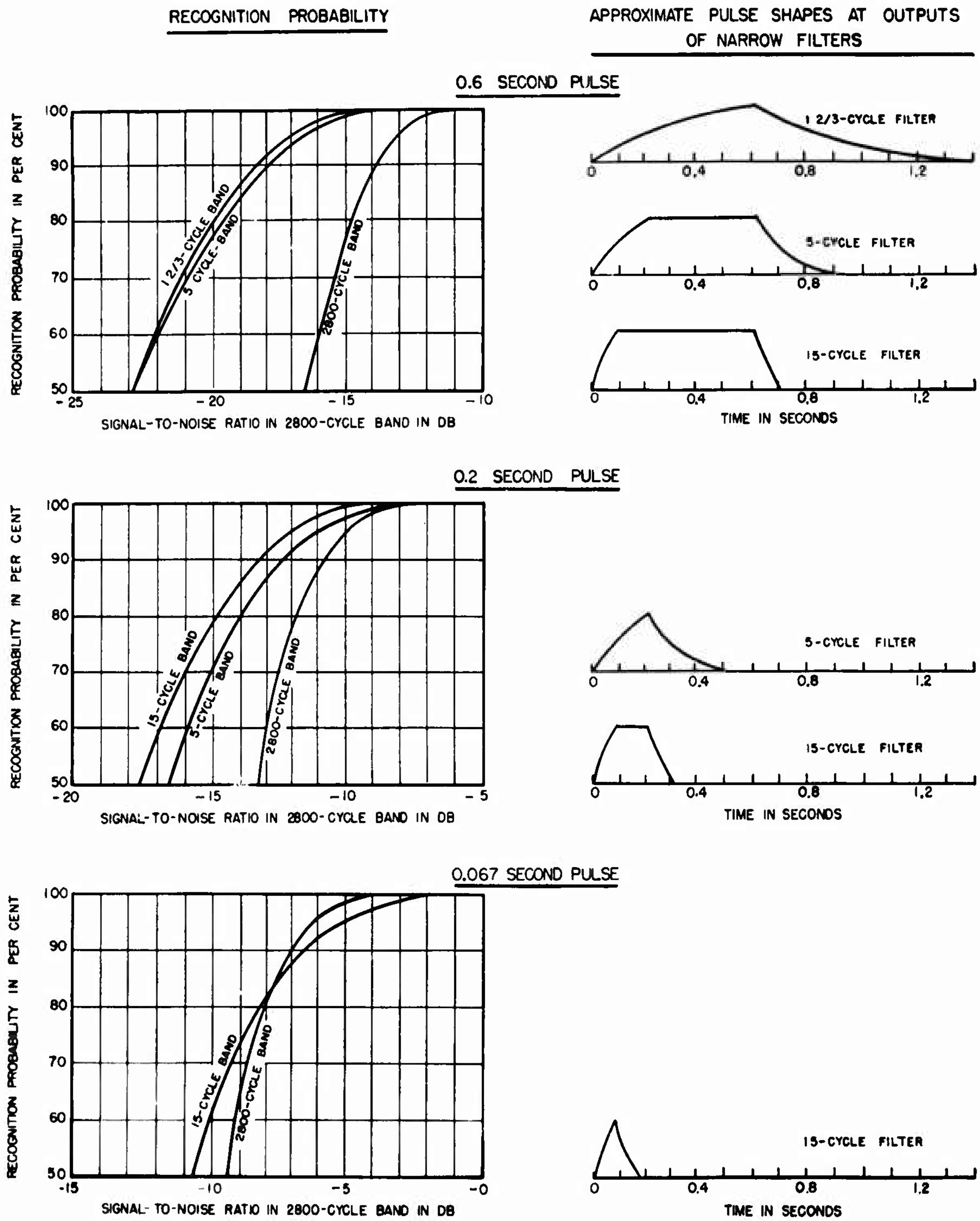


FIGURE 1. Effects of pulse length and band width on the detectability of pulses masked by noise.

this question should be answered in the negative. Thus, the peak intensity of a pulse primaudible at 740 cycles, at which frequency the critical band is about 38 cycles wide, should be equal to the intensity of the noise in that same critical band, in other words, equal to 38/2,800 of the uniformly distributed energy in the 2,800-cycle noise band. Hence, the overall signal-to-overall noise ratio at primaudibility should be -18.9 decibels (10 log 38/2,800). The observed values of the primaudible pulse-to-noise ratios and the apparent loss of audibility in decibels (which is equivalent to an apparent widening of the critical bands) relative to the primaudible sustained tone-to-noise ratio, are presented in Table 1 for the 3 pulse lengths

TABLE 1. Audibility loss relative to critical-band value.

Pulse length in milliseconds	600	200	67
Observed pulse-to-noise ratio in db	−16.6	−13.3	−9.3
Loss relative to −18.9 db in db	2.3	5.6	9.6

studied (see Figure 1). The same degrees of loss relative to a sustained tone are represented graphically in Figure 8, discussed at length in Section 8.4.4, where the filled-in triangles refer to the primaudible signal-to-noise ratios for the three pulse lengths named here, and the level of the horizontal line indicates the signal-to-noise ratio required for a very long signal.

Although the preceding analysis definitely establishes the degree of loss in audibility due to diminished signal length (and this loss is a significant quantity in practice), it does not reliably indicate whether the loss should be assigned to change in critical band width as a function of signal length, or to other factors. It has already been mentioned that two other relevant factors may be expected to influence the recognition of short pulses: the smaller stimulation produced by a short sound (which should, in the absence of a masking background, affect only those signals that are less than 250 milliseconds long), and the smaller likelihood that a short signal will be presented during a favorable low intensity interval in the noise. The latter factor (which will be designated the *sampling effect* in the following discussion) obviously depends on the time pattern and peak factor of the noise and may perhaps be profitably studied by the methods discussed in Chapter 9 in connection with reverberation backgrounds. On the basis of Figure 8 it would seem that the effects of these two factors may become significant for an individual pulse when its duration is less than 1,000 milliseconds, at least for thermal noise backgrounds.

It can be shown that any change of critical band width that may occur with changing signal length does not contribute significantly to the observed loss of audibility. That this is substantially the case, even for pulse lengths as short as 40 milliseconds, is indicated partly by the discussion of Figure 8 given in Section 8.4.4 and partly by the following examination of the data shown in Figure 1.

Consider any one pulse length τ, first, when the presented background extends over the full 2,800-cycle band, and secondly, when the presented background is transmitted by one of the narrow filters; then note the relative performance in the two cases. This procedure tends to eliminate all effects of pulse length associated with the ear's build-up time as such, although, of course, the build-up time of the test apparatus and therefore the physical nature of the stimuli are modified by insertion of the filters (these residual effects are discussed later). Similarly, it tends to eliminate all influence of pulse length associated with the sampling effect although, here again, changes of system band width complicate the situation by altering the envelopes of the pulses and the noise peaks (these effects are also discussed later).

It may be assumed that the difference in audibility observed with and without the narrow filter is not due to effects of pulse length, since that is the same in both cases. If the masking principles discussed in Chapter 2 are valid here, then the improvement in performance resulting from use of the filter is due to the fact that its admittance band B is narrower than the critical band, of width Δf, centered at the same frequency (at 740 cycles, Δf equals 38 cycles). In other words, use of the narrow filter reduces, in the ratio $B/\Delta f$, the transmitted intensity of the masking noise, that is, of those components in the noise background which are responsible for masking. This reduction of masking power should improve performance by

10 log $(B/\Delta f)$ decibels. When B is less than Δf, this quantity is negative; in other words, the signal-to-noise ratio needed for primaudibility is diminished by this number of decibels. Its absolute magnitude is entered in Table 2 as the "calculated improvement." When B is greater than Δf, the ear receives no assistance from the filter in discriminating between the frequency compositions of the pulse and the noise, and the preceding relation ceases to express anything significant for the process of auditory perception. Thus, when B is less than Δf but not less than $2/\tau$, performance should be improved by the indicated amount through use of the narrow filters, since the intensity of those noise components which are responsible for masking decreases as $B/\Delta f$, whereas the dominant components in the spectrum of the pulse are not attenuated. Before proceeding to examine the experimental data in detail, it should be noted again that the present analysis seeks to make a denial rather than an assertion; it is concerned with showing that the inferior audibility of tonal pulses as compared with continuous tones is not caused primarily by diminished pitch discrimination, that is, by a widening of the critical band with decreasing pulse length.

The quantities B and τ listed in Table 2, as well as the values of their products, refer to the data in Figure 1; in other words, the "observed improvements" represent the differences, for 50 per cent recognition probabilities and the stated pulse lengths, between the overall signal-to-overall noise ratios required for primaudibility with and without the narrow filters. The method of finding the "calculated improvement" has already been explained. The differences between calculated and observed improvements (the relative success of the preceding analysis in predicting improvement) are listed, for reasons stated below, under the heading "total distortion loss". Two cases are distinguished here, depending on whether $B\tau$ is less or greater than 2.

Table 2. Effects of band width and pulse length.

τ in seconds	B in cycles	Improvement in db		Total distortion loss in db		Loss due to pulse distortion in db
		Calculated	Observed	$B\tau = 1$	$B\tau = 3$	
0.067	15	4.0	0.7	3.3	---	---
0.200	15	4.0	4.3	---	−0.3	3.6
0.200	5	8.8	2.7	6.1	---	3.5
0.600	5	8.8	6.2	---	2.6	---
0.600	$1\frac{2}{3}$	13.6	6.2	7.4	---	---

It should be borne in mind, when examining the differences between calculated and observed improvements, or total distortion losses, that the values listed in Table 2 are much less reliable than the general trends, for which several factors are responsible.

First, the observed improvements are themselves computed by taking differences (hence, already affected by two experimental errors) between signal-to-noise ratios obtained in tests in which the number and identity of observers were not always the same and in which the sensation levels of the 50 per cent signals were quite different.[a]

Secondly, in several cases, a short extrapolation has been made in order to find the signal-to-noise ratio corresponding to 50 per cent detection probability.

Finally, the narrow filters used in these tests have a frequency response which is a typical resonance curve; the band widths of the equiva-

[a] It was observed during these tests that the gain setting adopted by the listeners was determined primarily by the level of the presented background. It may be inferred, therefore, that the gain settings generally corresponded to a loudness level of between 60 and 70 phons, that is, to a higher critical-band sensation level in the case of the filtered sound than in the unfiltered. As indicated in Section 9.2.2, this factor may influence the results, though probably by no more than 2 to 3 decibels.

lent rectangular filters assumed in deriving the computed improvements were obtained by finding the number of cycles included between points 6 decibels down from maximum response, which is only an approximation, though probably a fairly good one. Another approximation which is frequently used involves the −3-decibel points. Measurements made during these tests, of the fractional energy in the 2,800-cycle band transmitted by each of the narrow filters, give somewhat better agreement with the −3-decibel definition. Which is the most significant procedure to use, from the standpoint of the auditory process, is not clear; hence the values of the nominal band widths used in this section are the ones apparently preferred by the authors of the original report.[1] The reason for their preference is not specified, and the choice may well have been based on nonauditory considerations exclusively.

The trends, then, are as follows. When distortion is small ($B_T = 3$), the observed improvement is very nearly equal to the calculated, which confirms the assumption that critical band width is not substantially different for pulses and sustained tones. When distortion is significant ($B_T = 1$), the observed improvement is comparatively small. The difference between observed and calculated improvements is relatively large, in fact, a 5-cycle filter gives less observed improvement in the case of a 200-millisecond pulse than is obtained with a 15-cycle filter, and the magnitude of the disagreement between observed and calculated values increases as the band width of the filter is diminished, presumably because the resemblance is thereby increased between both the pitch and time-amplitude pattern of the transmitted signal and the transmitted background. In other words, constrictive filtering is injurious because it distorts the background as well as the signal (hence, the term "total distortion loss"); it may be, therefore, that the somewhat poorer agreement between observed and calculated improvements, for the 600-millisecond pulse, is partly due to the effect of the 5-cycle filter on the background.

The last column of Table 2 represents an estimate of the deterioration which may be expected from distortion of the pulse envelope alone. The estimate is in good agreement with the data for rounded pulses shown in Figure 8 and is of interest quite apart from constrictive filters, since the envelopes of real echoes may be distorted during transmission or reflection. This segregation has been made by noting that the observed improvement for a 67-millisecond pulse and a 15-cycle filter is 3.6 decibels less than obtained for a 200-millisecond pulse and a 15-cycle filter. Since the filter width is identical in both cases, its effect in distorting background is the same. Hence the loss of 3.6 decibels must be due to the fact that pulse distortion is greater for the 67-millisecond pulse ($B_T = 1$) than for the 200-millisecond pulse ($B_T = 3$), as indicated by the pulse shapes shown in Figure 1, and similarly for the pulses 200 and 600 milliseconds long in the case of the 5-cycle filter. It should be noted that the observed improvements used in computing the last column of Table 2 are themselves independent of such factors as build-up time of the ear and sampling, because of the way in which they were derived; hence, the major effect of pulse length reflected in the last column is distortion of the signal envelope.

It follows from the preceding discussion that narrow filters of optimum width (probably $B_T = 2$) may improve the performance of echo-ranging gear in the field under certain conditions. Unfortunately, however, these conditions are not compatible with the general requirements. Thus, narrow filters help significantly (bearing in mind that an improvement of 12 decibels in the detection process will double the maximum echo range only in the absence of severe thermal gradients and attenuation) when the pulse length is 600 milliseconds or more, and such long pulses would have the additional attraction of being more audible because of the diminished effects of sampling and auditory build-up time. Long pulses are not practical, however, because they would increase the search time, offer less security, increase the level of reverberation and would probably give somewhat less accurate auditory range determinations. Background admitted by filters less than some 200 cycles in width has an unpleasantly tonal sound, and, even with the standard 1- to 2-kilocycle band width of current echo-ranging gear, sonar operators complain of growing "ping happy" on long cruises. Furthermore, the use of filters less than about 400 cycles in width risks loss of doppler-shifted echoes (see Chapter 10). When noise is limiting, the existence of doppler has relatively little effect on the audibility of echoes (see Figure 14, also equation (4) in Chapter 10) and means are available by which the amount of doppler in the presented echo can be reduced nearly to zero (see Chapter 10), but when reverberation is limiting, the existence of doppler tends to offset the masking effect of background (see Section 10.2.2). Furthermore, doppler gives tactically valuable information; hence, attempts have been made to increase the amount of the doppler shift in the presented sounds. Finally, Figure 1 indicates that the relative advantage afforded by the narrow filters at the 50 per cent level does not persist at the higher signal-to-noise ratios; in other words, the rate at which detection improves with increasing signal-to-noise ratio is greater for rectangular pulses, and n is reduced,[b] as a result of the distortion associated with the narrow filters, from 3.6 (without filters) to between 2.0 and 2.7 (with filters).

These values of n are comparable to values obtained in other tests with pulse signals which are described later, and, it will be noted, are much the same as the values obtained from tests with ship signals, which were discussed in Chapter 4. In other words, the relation between perception probability and signal-to-noise ratio does not appear to be strongly dependent on the band width of the masking background (such as noise and reverberation), the method of test administration, sensation level in the optimal band, or duration and time pattern of the primaudible component.

The method of administration and scoring used in the tests under discussion provides information on the relative frequencies with which actual signals are missed and nonexistent signals reported under various conditions. The available data are plotted in Figure 2, where both types of errors are expressed as fractions of the total number of signals per test,[c] and where both were found to increase progressively with diminishing signal-to-noise ratio. In the case of the CW pulses, there is a clear tendency for E_+/E_- to be larger in the tests involving constrictive filters ($B_T=1$) than in the others. Since constrictive filtering has less immediate interest, the line in this figure has been drawn to follow the remaining points. It will be observed that the ratio of E_+ to E_-,

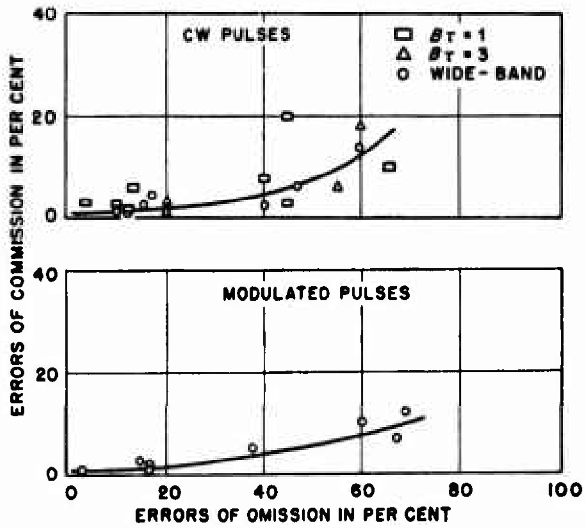


FIGURE 2. Relative frequency of errors of omission and commission.

which may be used to specify the tendency to give false reports, does not exceed 0.1 when the detection probability is better than 60 per cent (fewer than 40 per cent errors of omission), and that this is about equally true of undistorted CW pulses and of modulated pulses. When the detection probability is less than 60 per cent (more than 40 per cent errors of omission), the relative number of false reports becomes significantly greater for CW than for

[b] For a definition of n, see Section 4.1.4. If the "spread" of the transition curve is defined as the signal increase in decibels between 20 and 80 per cent recognition probability, then n equals 12 divided by the spread.

[c] The number of commissive errors per signal may be a function of the interval between signals and of the overall length of the testing period. In other words, it would be useful to know whether tests patterned after the field situation (few signals, long listening period) would increase the relative number of false reports, or failures to report, or both. Such a test, for the case of ship signals, is discussed in Section 4.2.6, but nothing similar has yet been described for pulse signals.

modulated pulses.[d] When a telegraph key was substituted for the observers during one part of this test program, the automatic detector was found to give a larger $E_{\cdot}/E_{-}$ ratio than did the ear for the same test conditions.

In organizing the data relating to masked thresholds of pulses, either of two standard reference bands may be used to define the prim-audible signal-to-noise ratio: (1) the 1,000-cycle reference band which is typical of many current echo-ranging installations (and this definition of the *RD* applies to the ordinates at the right of Figure 8), or (2) the somewhat more fundamental critical band, which may be taken as very nearly 50 cycles wide for all CW echoes heterodyned to frequencies between 0.1 and 2.5 kilocycles (and this definition of the *RD* applies to the ordinates at the left of the same figure). When the reference band width is not specified in the ensuing discussion, the critical band *RD* is to be understood. This has the advantage that all recognition differentials are positive — or zero, in the case of a very long pulse or sustained tone. Furthermore, it focuses attention on those components in the background which are most directly responsible for masking. To convert from one reference band to the other, in the case of a background with uniform composition (which is usually true of the field, as well as the laboratory situation), note that the *RD* for a 1,000-cycle band is 13 decibels less than for a 50-cycle band (10 log (50/1,000) = −13 db). Since critical band width varies somewhat with frequency, the use of the 50-cycle approximation between 0.1 and 2.5 kilocycles gives results slightly different from what would be obtained with the correct critical band width, but such differences are usually less than 1 decibel. Thus, the values designated as the loss relative to a sustained tone, which are given in the first table of the present section, are the critical band recognition differentials for rectangular CW pulses of the indicated durations.

When the duration of such pulses is less than $2/\Delta f$ seconds, where Δf is the critical band width, the essential spectrum of the pulse extends beyond the confines of a single critical band. If the stimulation of two adjacent critical bands is no more effective than stimulation of either alone, it might be expected that the progressive broadening of the signal spectrum with diminishing pulse length would produce a more rapid deterioration, in the audibility of CW pulses masked by wide band noise, than is observed for pulse lengths exceeding $2/\Delta f$ seconds. For $f = 50$ cycles, the change should set in at a pulse length of 40 milliseconds (2/50 = 0.04 second).

Two observations discussed in Sections 8.4.2 and 8.4.5 lend support to these remarks. First, the rate at which the audibility of 800-cycle CW pulses, less than about 50 milliseconds in duration, deteriorates is greater than that for longer pulses. Secondly, 6-kilocycle CW pulses with durations of 10 milliseconds or less are found to be only 1 to 2 decibels less audible than 1-kilocycle pulses of equal duration, although, on the basis of the critical band widths at these two frequencies, one would expect the 1-kilocycle and 6-kilocycles pulses to differ in audibility by nearly 8 decibels [10 log (300/50) = 7.8 db]. The discrepancy of some 6 decibels is of the expected magnitude, at least for a 10-millisecond pulse, as, in this case, the essential band is about 200 cycles ($2/\tau$) wide. Therefore, when the intrinsic frequency is about 1 kilocycle, the pulse energy is distributed among four contiguous critical bands, each about 50 cycles wide. Although the energy is unequally distributed among these four bands, equality of distribution may be assumed as a first approximation. If detection occurs when the signal-to-noise ratio assumes the same value, *R*, either in one of these four 50-cycle critical bands or in the single 300-cycle critical band centered near 6 kilocycles, then the 6-kilocycle pulse should be four times more audible than the 1-kilocycle pulse, and 10 log 4 = 6 db.

There is also another set of observations, described in Section 2.2.2, in connection with the apparent loudness of pulses, which seems to have some bearing on the effects that occur when the essential spectrum of a tonal pulse extends beyond the limits of the critical band stimulated by the intrinsic frequency of the

[d] This observation does not necessarily mean that modulated pulses are to be preferred in practice, since, as shown later, they may be less detectable than CW pulses. The point here is merely that modulated pulses offer the listener an additional cue; he is, therefore, better able to classify some of the "doubtful" perceptions.

pulse. In that section, the discussion was concerned largely with the subjective loudness of the pulses; in the present section, the data of

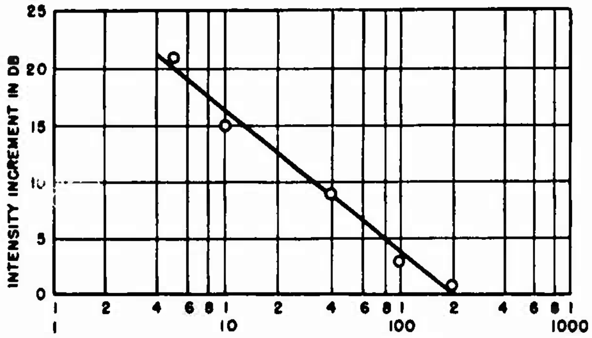


FIGURE 3. Increments needed to maintain 125-cycle pulses at a loudness level of 60 phons.

Figures 12, 13, and 14 in Chapter 2, are replotted in terms of the intensity levels (actually, the sensation levels) of the sounds, in order to simplify comparison with the masking data shown in Figures 1 and 8. The replotted data are shown in Figures 3, 4, and 5, which were obtained by reading the points of intersection between the curves in Chapter 2 and either of two horizontal guide lines (the 60-decibel line in one case, and the 20-decibel line in the other). For each of these two cases, the horizontal distance between the dotted line and the above-mentioned point of intersection represents the number of decibels (called "the in-

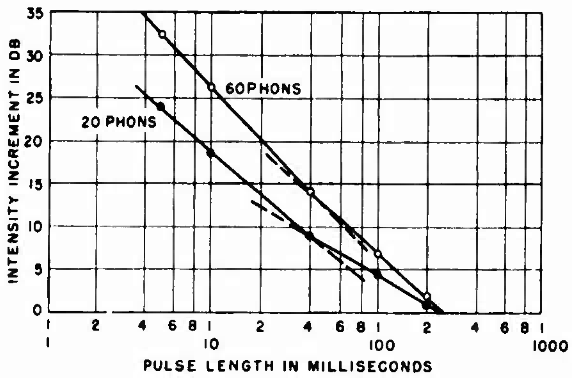


FIGURE 4. Increments needed to maintain a 1-kilocycle pulse at the indicated loudness levels.

crement" in what follows) by which the level of a pulse of the indicated duration must be increased in order for it to produce a subjective loudness equal to that of a particular sustained tone. Clearly, this sustained tone has a sensation level of 20 decibels in one case and of 60 decibels in the other. This procedure was carried through for each of the three frequencies shown in Figures 12, 13, and 14 in Chapter 2. In order that the increments so obtained represent the number of decibels increase required for the tonal pulse in question, and not merely that of the equally loud 1-kilocycle tone, it is necessary to express the data of Figures 12, 13, and 14 in Chapter 2, in terms of sensation, rather than loudness level. These quantities are equal (by definition) at 1 kilocycle and essen-

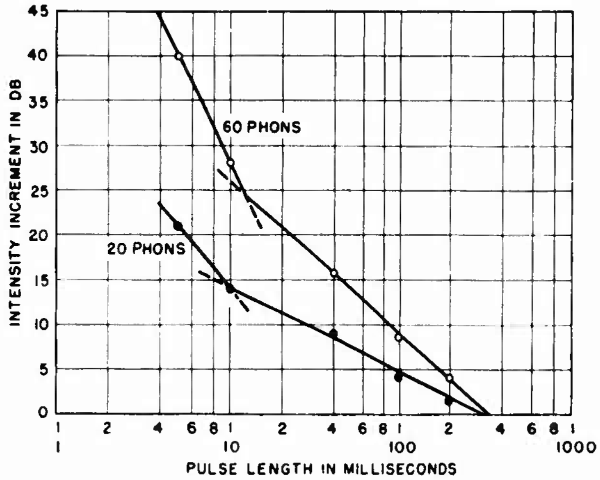


FIGURE 5. Increments needed to maintain a 5.65-kilocycle pulse at the indicated loudness levels.

tially equal at 5.65 kilocycles; they are quite different for a 125-cycle tone. The translation may be made by means of Figure 11 in Chapter 2.

Owing to the paucity of the data and the curvature shown by the curves in Figures 12, 13, and 14 in Chapter 2, in the neighborhood of 20 phons, the results for the lower loudness level are less reliable than those for the higher. For this same reason, the curve for 20 phons has not been included for the case of 125 cycles in Figure 3.

It will be clear from the preceding description that the increments shown in Figures 3, 4, and 5 represent a loss in loudness due to auditory build-up time. They are plotted in such a way as to be readily comparable with the loss of detectability (due to all causes and shown in Figures 8 and 14) which is suffered by a pulse, in the presence of masking noise, when its duration is reduced. In other words, the indicated increments are the amounts by which the intensity levels of pulses with various durations must be augmented so that each pulse will produce the same loudness (and hence, the same stimulation) as a sustained tone of specified loudness.

Examination of the increment graphs shows that the points can best be fitted by a single straight line for the 125-cycle tone, but that two straight lines give a better fit for the 1- and 5.65-kilocycle tones. Where two lines are needed, they intersect at a pulse length very nearly equal to $2/\Delta f$. Since Δf is 260 cycles at 5.65 kilocycles, and 50 cycles at 1 kilocycle, the discontinuities should occur at pulse lengths of 8 and 40 milliseconds, respectively (2/260 and 2/50); the observed breaks come at 10 and 40 milliseconds. Adopting the criterion that the break should occur at a pulse length of $2/\Delta f$ seconds is equivalent to assuming that (1) any segment of the basilar membrane, with dimensions corresponding to a critical band, constitutes a resonance element of band width Δf; and (2) the response of such a resonance element suffers when $\Delta f \tau$ is less than 2. Thus, when τ is diminished, since Δf is approximately independent of τ, the maximum in the basilar vibration pattern ceases to be as well localized as for longer tones, and the pitch and loudness functions show a more or less abrupt change; in other words, progressively larger numbers of its "channels" begin to function cooperatively, and the basilar membrane tends to assume some of the characteristics of a broadly tuned system. The 125-cycle graph shows no discontinuity, presumably because the low sonic frequencies already stimulate practically the entire basilar membrane, even at low sensation levels; in other words, pitch perception in this region differs from that at higher frequencies. A precautionary note should be added at this point: the increment graphs are based on a relatively small amount of data; they provide an indication rather than a proof. Moreover Figures 3, 4, and 5 have been obtained from the smooth curves drawn in Figures 12, 13, and 14 in Chapter 2; the scatter of the observed points in the latter three figures leaves the conclusions of the present discussion open to some doubt. These tentative conclusions are, however, in agreement with the trend of the recognition differentials for short pulses discussed in Section 8.4.4.

In connection with the response time of the ear, it should be noted that some sense of pitch remains even for relatively short stimuli. Thus, the value of the ear's response time depends, to some extent, on the task set and the criterion used to rate its performance. For example, loudness begins to deteriorate at about 250 milliseconds, presumably due to summation effects in the acoustic nerve. The pitch function seems to show a discontinuity at $2/\Delta f$ seconds, but pitch perception is possible (in the middle range of frequencies) down to 10 milliseconds for rectangular pulses and to 3 milliseconds for rounded pulses, and these limits seem to be set by the response time of the overall mechanical structure of the ear. Finally, judgments of phase differences at the two ears permit subdivision of the 1-millisecond interval between successive impulses conducted by a nerve fiber, so that sounds may be localized even when the difference in their arrival times is of the order of 10 microseconds. Each of these durations is significant for a specific problem, but no one of them is *the* response time of the ear.

Sensation levels of 20 and 60 decibels have been adopted in constructing the increment graphs because these values are probably fairly close to the upper and lower limits encountered in standard echo ranging practice, and when receiver gain is set for a comfortable listening level. The slopes of the increment graphs depend on sensation level, the rate of deterioration with falling pulse length being greater at the higher sensation levels, although the intercepts are in all cases of the order of 250 milliseconds. It cannot be assumed, however, that better recognition differentials will be obtained when the gain is set for a low listening level, since obviously several factors besides auditory build-up time combine to give the observed dependence of primaudibility on pulse length. The fact, alone, that recognition differentials begin to deteriorate when the signal is reduced below about 1,000, instead of 250 milliseconds, is sufficient indication that this is so. Furthermore, the transient click which contributes to the broadened spectrum of a short pulse, and which is associated with loss of tonality, and also with the augmented loss of loudness for pulses less than $2/\Delta f$ second in duration, is itself a useful cue when a pulse must be detected in the presence of masking noise.[e] Assistance from this factor will, of course, be greater for the longer pulses (in which the click is connected with and calls attention to a signal whose duration exceeds that of the random noise peaks). When the signal length is very short, approaching the duration of the noise peaks, relative squareness of the signal envelope ceases to constitute a very significant cue. In the latter case, the relative intensities of signal "pops" and noise "pops" tends to be the factor of major importance, and the results obtained for pulses of different frequency and envelope tend to be much the same.

Intensity alone, however, is not an altogether satisfactory basis for echo detection. Consider a 1-millisecond pulse; according to Figure 8, such a pulse becomes primaudible (in a short laboratory test, and when the observers know that signals are presented fairly frequently) when its level is about 12 decibels above the mean level of the noise in a 1,000-cycle presentation band. In the field situation random noise bursts, not readily distinguishable from 1-millisecond echoes on any ground but amplitude, are commonly encountered in the masking background. This is due to the fact that pitch perception is practically nonexistent for 1-millisecond pulses; in fact, if the intrinsic frequency of the heterodyned pulse is below 1 kilocycle, less than one cycle will be completed. In addition, the duration of noise peaks transmitted through a band width of 1,000 cycles is, like that of the signal pulses, about 2 milliseconds ($2/B = 2/1{,}000 = 2$ milliseconds); in other words, for the system assumed above, noise peaks and signals will have durations of about 2×10^{-3} second. Consequently, the average number of noise peaks per second will be about 500, or $[1/(2 \times 10^{-3})]$. If the time-amplitude pattern of the actual background resembles that typifying the class of sounds grouped under the head of thermal noise, then the probability that a random peak will have 16 times ($10 \log 16 = 12$ db) the average intensity is e^{-16} or about 10^{-7}.[f] Hence, the total number of noise peaks per second with intensities equal to or greater than 16 times the mean noise level is 500×10^{-7}, or 0.5×10^{-4}. During a watch of one hour, therefore, at least one noise peak of this amplitude will occur. Notice, however, that peaks 10 decibels above the average level will occur at a rate of $500 \times e^{-10} = 500 \times 10^{-4.3} = 0.025$ per second, or better than 1 per minute. Thus, a listener may fail to notice a 1-millisecond pulse when it is only 12 decibels above the noise in a 1-kilocycle band, and when he is not expecting such a pulse. This fact may have important applications in the use of submarine echo-ranging gear and acoustic fathometers, which are often required to yield a detectable echo with only a few pulses. It will also be difficult for the operator of antisubmarine echo-ranging gear to be sure that he has detected a single very short echo. Sending a sequence of 2 or 3 short pulses and giving them a "signature" or rhythm which is not likely to characterize successive noise peaks may aid positive identification. Some observations were made in a series of field tests regarding the effect of signature[3] in which two surface vessels equipped with echo-ranging gear participated. The operator on the target vessel often failed to distinguish between the random pops in the water noise and short pulses arriving from the other vessel, even when such pulses were fairly intense. Similarly, the operator on the echo ranging vessel was not always able to classify these short echoes with certainty although he had the advantage of knowing whether or not a pulse had been put into the water. Overheard signals, as well as returned echoes could be very much more readily classified under given noise conditions when a pair of pulses separated by between 50 and 100 milliseconds were used instead of a single pulse. Such a pair of pulses was recognized as a distinctive "dit-dit."

In general, the slopes of the increment curves in Figures 3, 4, and 5 are in the neighborhood of 15 decibels per decade; that is, the intensity for constant sensation level, in the absence of masking, is roughly inversely proportional to the 3/2 power of the pulse length. This rela-

[e] It is, in fact, by responding to the spectra of brief pulses, rather than to their envelopes as such, that the ear is able to distinguish between pulses of the same duration and intrinsic frequency but different shapes.

[f] This follows from the fact that the distribution of amplitudes for random noise obeys the Rayleigh law (see Chapter 7 and reference 2 in this Chapter).

tionship is a convenient way to summarize the data; it cannot at the present time be considered to have any fundamental significance, since the precise meaning of sensation level in terms of auditory mechanics is not known. In fact, the available information indicates that the magnitude of the smallest stimulus which will produce discharge of a nerve fiber is an exponential function of stimulus duration. It is not certain to what extent this behavior depends on general fatigue, nor how much it varies among normal, unfatigued observers. But it seems likely that its operation would tend to increase the range of variability among recognition differentials obtained from a representative group of subjects listening to pulses, so that the variability among individuals would be greater than is usually observed when sustained signals are used.

It is interesting to compare the slopes of the increment graphs with the rates at which intensity discrimination deteriorates for amplitude-modulated signals (see Figure 16 in Chapter 2). In this latter case, it will be recalled, the amplitude of tones was modulated sinusoidally at various modulation rates, and the intensity limen was observed to increase progressively for modulation rates in excess of 3 cycles per second. Detection of a loudness change under these conditions is equivalent to recognizing a signal masked by a fixed-level background (see the discussion of Figure 54 in Chapter 4), thereby differing in one fundamental respect from detection in the presence of a fluctuating, thermal noise background. The signal used in the intensity discrimination tests may be regarded as consisting of a string of rounded pulses, each pulse being one modulation cycle in length. This suggests plotting the deterioration in loudness discrimination against the duration of the modulation cycle, for purposes of comparison with the increment and the *RD* graphs. This replot is shown in Figure 6, and it will be observed from this figure that the rate at which intensity discrimination deteriorates is about 10 decibels per decade. In addition, the intercept comes at about 300 milliseconds, which agrees with the behavior of the increment graph but is in contrast to the 1,000-millisecond intercept shown on the *RD* graphs. The difference between the positions of the intercepts for the *RD* and the intensity discrimination data is perhaps due in part to the fact that the effective background component in the intensity discrimination tests had a fixed

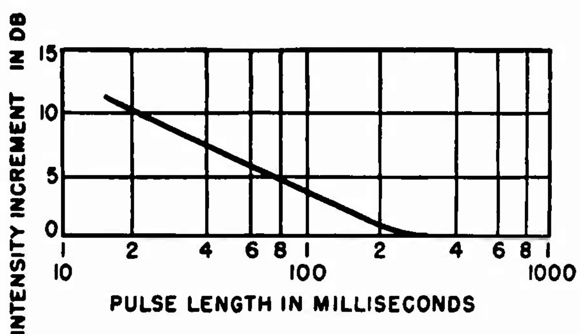


FIGURE 6. Increments needed to maintain detectability for strings of rounded pulses.

level, whereas the masking background in the echo tests consisted of fluctuating thermal noise; this effect is discussed more fully in Section 8.5.

From the data plotted in Figures 3, 4, 5, and 6, it is evident that the intensity increment for constant loudness varies between 10 and 20 decibels per decade of pulse length, the precise value depending on the experimental conditions. This relationship holds only for durations less than 250 milliseconds. For longer durations, full subjective loudness occurs, determined only by the intensity level and frequency of the sound. For sounds of less than 250 milliseconds duration, the ear's response depends more upon the total energy in the pulse (power times duration) than it does on the power alone.

It is sometimes assumed that for short pulses the response of the ear is directly proportional to the total energy. Figures 3 through 6 show that in some cases the response seems to vary more rapidly with pulse lengths than would be expected from this simple relationship. The available data indicate that, for constant loudness, the intensity of the pulse should be increased more nearly as the inverse $\frac{3}{2}$ power of the pulse length than as the inverse first power. In view of the lack of accurate data, however, it is possible that the inverse first power is more nearly the correct relationship and that for short pulses the ear's response is determined primarily by the energy of the pulse.

Various aspects of the auditory process have been reviewed and correlated in the foregoing pages in order to provide a certain amount of setting and perspective for the observations described in the remainder of this chapter. In general, recognition of CW pulses in the presence of noise depends on the spectrum and time pattern of the background, the band width and center frequency of the signal spectrum, the duration of the pulse, and the integrating time of the ear.

8.1.3 Detection of FM and AM Signals

Measured recognition differentials for each of three kinds of modulated signals masked by noise are shown by the open triangles in Figure 8. The width of the noise band as well as other details of procedure were the same as used in the tests with CW signals. The open symbols have been entered in Figure 8 for the sake of comparison with the points for rectangular CW pulses which are indicated by means of filled-in symbols. The line has been drawn to fit these filled-in symbols only.

The open triangle appearing in the graph at a pulse length of 200 milliseconds refers to a pair of 200-millisecond pulses (with an intrinsic frequency of 740 cycles) which were separated by an interval of 200 milliseconds. This corresponds to a square-wave amplitude modulation of the CW signal. Comparison with the *RD* for a single 200-millisecond pulse (shown by the filled-in triangle) indicates that a single repetition helps by about 3 decibels under these conditions. If the critical band *RD* has positive values for pulses between 250 and 1,000 milliseconds long solely because of the operation of the sampling effect, then the latter accounts for a loss in audibility of about 5 decibels for a 250-millisecond pulse (compare Figures 4 and 8). Since a single repetition offsets 3 decibels of this 5-decibel loss, it is clear that additional repetitions can provide little further improvement. This statement must be modified when the pulse repetition frequency exceeds 18 to 20 per second, in which case additional factors may enter into the situation (see Section 8.5). For such high repetition rates, however, the duration of the individual pulses must be diminished if the signal is to consist of a string of pulses rather than a sustained tone. In one set of tests (see Figure 18), repetition of a 1-millisecond pulse at a rate of 5.6 per second, continued for several seconds, appears to have helped by about 10 decibels with respect to the *RD* for a single pulse. Unfortunately, not enough data are now available to indicate whether this result is reliable.

It does not appear likely that the improvement observed when the double pulse was used arose from the integrating action of the auditory mechanism, because the major effects of the first of the two pulses are not likely to have persisted more than about 0.14 second after its termination (see also Section 8.2). Nevertheless it is interesting to compare the audibility of this repeated 200-millisecond pulse with that of the 600-millisecond pulse, since both have equal overall duration. It will be seen from Figure 8 that performance obtained with the repeated signal is very nearly as good as that obtained with the 600-millisecond pulse. Use of the dit-dit signal would therefore give a slight advantage relative to the longer one when reverberation is limiting, because the strength of background would be reduced somewhat by abbreviating the transmission. In addition, it is probable that losses of echo intensity due to rapid changes in transmission conditions will not affect both members of such a pair of pulses equally, and in consequence the chance of detecting at least one of them would be improved. However, the use of a twin pulse is time consuming and is, in general, undesirable during general search and screening operations.

The open triangle plotted at 400 milliseconds refers to a 400-millisecond signal consisting of two successive tones (0.8 and 1 kilocycle), each of 200 milliseconds duration. Since the ear can very nearly achieve maximum response during each of the 200-millisecond subintervals, the *RD* for such a signal would be expected to, and does, approximate the value for a CW signal with equal overall duration and with an intrinsic frequency of 740 cycles; in other words, critical band width is nearly equal for all three frequencies.

The open triangle plotted at 110 milliseconds refers to a frequency-modulated pulse. This had a constant amplitude and its frequency was increased at a uniform rate from a value of 400 cycles at its beginning to a value of 2,500 cycles at its end. The recognition differentials for this pulse and for the pulse described in the preceding paragraph were computed from the ratio of signal intensity to intensity of noise in a 1-kilocycle reference band; that is, by adding 10 log 2,800 1,000, or 4.5 decibels, to the observed *RD* for the 2,800-cycle presentation band. This procedure would be a useful basis for comparison even if the critical-band concept could not be invoked; but there is no reason to suppose that auditory discrimination of pitch was involved in any fundamentally different way for the CW and for the two-frequency and multi-frequency pulses, as indicated later, and by the general consistency between the recognition differentials for the three kinds of pulses.

It will be observed that the FM pulse was about 2 decibels less audible than would be expected for a CW pulse of equal length. Field observations also indicate[3] that, in the presence of noise, an FM echo about 150 milliseconds long is approximately as audible as a CW echo of equal amplitude and duration. It follows that the *RD* plotted for the 110-millisecond FM pulse in Figure 8 is fairly reliable, and it is worth considering some of the factors which enter into this result, for the light it may shed on the auditory process involved, and also because FM pulses may have certain practical advantages when reverberation is limiting.

As to the reliability of the laboratory tests with FM pulses, it should be noted that in the experiments described in the present section, the masking noise was always passed through the 2,800-cycle filter, and, in some tests, through the narrow filters as well. The pulses were always filtered when the narrow noise backgrounds were used, but not in other tests. When the spectrum of a pulse extends beyond the limits of the noise spectrum, and thus outside the masked region, the results of listening tests may be misleading since, in the field, signal and noise must reach the ear over the same system. This factor was of no consequence in the CW test series, since the essential spectra of all CW pulses were less than 2,800 cycles in width. Furthermore, the general agreement between field and laboratory observations indicates that no substantial error resulted from failure to filter the FM pulse.

It is evident, first of all, that the results for FM pulses give very significant information on the nature of hearing. They strongly suggest that the build-up time of the ear is related to some post-cochlear stage of the auditory process. An FM pulse 110 milliseconds in length, sweeping over 2,100 cycles, takes only 3 milliseconds to sweep over a critical band 35 cycles in width. Thus, if each critical band had its own build-up time, this FM pulse would be expected to have a recognition differential comparable with that of a 3-millisecond CW pulse. Actually, the observed *RD* corresponds to a CW pulse of about 80 milliseconds. It may be inferred, therefore, that there is apparently an integrating center involved in the hearing process which can correlate events at widely separated points on the basilar membrane, summing these separate events into a single whole.

It may be possible to test this conclusion by studying the loudness of FM pulses as a function of overall duration and frequency sweep, although reliable loudness balances might be more difficult to make for FM than CW pulses.

As to the detailed auditory process involved, it should be noted first that, in the field and in the presence of masking, FM echoes are heard as chirps or glides, provided the echo is not too brief or the frequency sweep too extended. It seems reasonable, therefore, to consider the masking of FM pulses from the point of view applied to the problem of the sensed properties of glides as heard in the absence of masking.

In the case of a glide, the physical stimulus may be considered[4] as composed of a sequence of brief rounded pulses having slightly different frequencies and, of which, no more than two or three need be assumed to have substantial amplitudes simultaneously. Thus, the coexisting pulses combine to give a wave with constant amplitude and variable frequency. At the beginning and end of the glide, however, the component pulses have abrupt onsets or terminations, so that the FM wave must be considered as made up of large groups of frequencies at its termini, and these transients are heard as clicks without significant tonality. Between the clicks, the successive pulses stimulate essentially the same segments of the basilar membrane as respond to sustained tones with the intrinsic frequency of the pulse. The locus of the maximum amplitude of basilar disturbance associated with successive pulses moves smoothly from one pitch region to the next because of the way in which the phases and amplitudes of the different pulses are related. Furthermore, the sensed pitch moves up or down the scale in much the same manner as does the frequency of the objective stimulus. The relative absence of tonality at the ends of a rectangular FM pulse causes the perceived extent of the glide to be less than its actual extent; hence, it may be best to use rounded rather than rectangular FM pulses.

In terms of the line-busy effect, the ability to perceive the various frequencies in an FM signal depends on whether or not the maximum amplitude of the engendered disturbance, which moves along the basilar membrane, exceeds the effective amplitude of the noise tending to jam reception in each of the critical bands successively stimulated by the glide. The results shown in Figure 8, and those obtained in the field, imply that (for a 100- to 150-millisecond pulse, with a sweep of 1 to 2 kilocycles) the effective amplitude at the maximum of the moving disturbance is very nearly of the same magnitude as that produced in the case of a CW pulse whose intrinsic frequency coincides with the center frequency of the glide.

The relative audibilities of various FM pulses would be expected to depend on a number of factors. Thus, the rate of sweep should probably not be too great. To make this statement somewhat more quantitative, consider the FM pulse used in the tests under discussion. This was swept through 2,100 cycles in 0.110 second. The disturbance produced on the basilar membrane by such a stimulus probably has a fairly well localized maximum at any instant, but this maximum is very likely of appreciable amplitude over a segment corresponding to not less than the critical band width (about 50 cycles at the middle frequencies). Hence, a disturbance which sweeps over 2,100 cycles in 110 milliseconds will produce nearly maximum stimulation of a short segment of the membrane for a period of a few milliseconds; and this duration begins to approach the time threshold of pitch perception. This effect might be modified to some extent by increasing the listening level, which should increase the width of the maximum and thereby increase the length of time during which any segment of the membrane is stimulated by the traveling disturbance.

Furthermore, the critical bands stimulated by a pulse which is swept from a frequency of 400 cycles to one of 2,500 cycles are not equally wide, and their widths (or relative amounts of masking produced by backgrounds) at the lower and upper limits of sweep differ by a factor of nearly 3, or about 5 decibels. Thus, variations in critical band width alone seem sufficient to account for the 2-decibel difference in the audibilities of the approximately 800-cycle CW pulses and the FM pulse referred to in Figure 8. It is reasonable to infer that the recognition differential for an FM pulse against a noise background may be smallest (most favorable) if the pulse is restricted to the frequency region in which the critical band width is less than 50 cycles; in other words, in the region between 100 and 1,500 cycles.

When the duration of a rectangular FM pulse is reduced below about 30 milliseconds, the interval during which the ear senses the initial and terminal clicks becomes a very large fraction of the overall pulse duration. In addition, if the sweep is large, the time which the signal takes to sweep over a critical band width becomes so short that pitch perception disappears. These two effects, together with the possibly related build-up time of the ear as a whole, may be expected to impair the audibility of FM pulses.

Field tests indicate the extent of this impairment.[3] It was observed that a 10-millisecond FM pulse was very much less audible in the presence of noise than a 10-millisecond CW pulse, unless the extent of sweep was small. The sweep should probably extend over at least one or two critical bands, however; otherwise, the glide becomes difficult to perceive, and the audibility of the glide itself may be a valuable means of distinguishing between real echoes and misleading simulations. Thus, during the laboratory tests with FM signals, it was "the almost unanimous opinion of the observers that frequency modulated signals were more readily detected than single-frequency signals." [1] The data taken during the same tests, however, show that the recognition differentials for FM pulses were actually less favorable than those for CW pulses of equal duration. In confirmation of this result, Figure 2 implies, not that the FM signals were more audible than the CW, but that they were easier to distinguish from false cues when such signals could be heard; that is, E_1/E_2 is smaller for primaudible signals when they have a distinctive signature. This may have a further advantage in practice by reducing uncertainty, tension, and fatigue. It should be noted, in this connection, that the slopes and shapes of the transition curves obtained with modulated signals were in all respects similar to those obtained in tests with CW pulses.

By way of conclusion, it seems worth suggesting that further study may profitably be given to the masking of rectangular and rounded FM pulses by noise and reverberation backgrounds, and as a function of the modulation parameters: rate, range, direction, and type of sweep (such as linear and sinusoidal).

8.2 CUDWR-USRL TESTS

In these laboratory tests, the results of which have been informally communicated, masking was studied for rectangular CW pulses of 16 milliseconds duration, and with intrinsic frequencies of 400, 800, and 1,600 cycles. One test was also conducted with a 1,000-millisecond pulse of 800 cycles frequency. The masking thermal noise was 300 cycles wide and centered at the pulse frequency. In general the apparatus used was similar to that described in the preceding section. The test sounds were presented to groups of between 6 and 10 observers by means of a high-quality loudspeaker and at a comfortable listening level. The test sequences consisted of groups of 30 listening intervals, each interval containing either a signal or a blank. In any sequence of 30 presentations, the noise level was fixed, and the signals were randomly and equally distributed between 6 predetermined levels (not including the blank level) which covered a range of 12 decibels in 2-decibel steps. This 12-decibel range was selected to include detection probabilities from nearly zero to nearly 100 per cent. The procedure described here (randomized pulse levels) differs from that discussed in the preceding section, since each test sequence in the former study featured only one signal-to-noise ratio. The fact that the results of the two types of test are so consistent (see Figure 8) implies that this difference in procedure is not very important. Listening intervals were about 3 seconds long, and were separated by silent intervals of the same length, during which the listeners reported their judgments of signal audibility. Rest periods were taken between the test sequences.

Changes of overall gain did not affect performance significantly, which indicates that, at the listening levels used, masking was produced exclusively by the thermal noise background. However, the level of room noise was very much lower than would be encountered on the bridge of an antisubmarine vessel, and it may be that the distracting and fatiguing effects of loud, environmental sounds are more important in the field than are the masking effects of those sounds.

The observed, primaudible signal-to-noise ratios for the 300-cycle noise bands were converted to the equivalent critical-band RD (using the average 50-cycle width) by adding 10 log 300/50, which amounts to 7.8 decibels. The critical-band recognition differentials obtained in this way are entered in Figure 8 as filled-in squares. The justification for use of the critical-band concept in the case of pulse signals has already been discussed. The average critical band width of 50 cycles was adopted because no significant differences in performance were observed for the three pulse frequencies used, although it should be added that the observers found the lower frequencies pleasanter than the high, because the 300-cycle bands of noise were less shrill when they were centered at the lower frequencies and, in fact, the observers felt that they were making finer auditory discriminations in the case of the 400-cycle pulse than for any of the others. Qualitatively, however, the 16-millisecond pulses appeared to have little definite tonality, irrespective of their intrinsic frequencies.

In evaluating recognition differentials, scores were based exclusively on the performance of observers who were guilty of no more than one error of commission per test sequence (6 blank intervals being presented in each sequence). In all tests, the frequency of commissive errors in the blank intervals did not exceed 7 per cent. From this, it may be inferred that the curve in Figure 2 does not continue to rise very steeply in the region where errors of omission exceed 60 per cent.

In agreement with other observations, the slopes of the transition curves obtained in the 16-millisecond tests give a value of n which lies between 2 and 3. For the 1,000-millisecond tests, the value of n was very nearly unity. However, the amount of data in the latter case is too small to warrant attaching any significance to this difference. Transition curves were obtained by finding the detection probability, at a given signal-to-noise ratio, for the entire group of observers. Transition curves were also drawn for the individual observers and a tabulation was made of the various signal-to-noise ratios at which each individual perceived 50 per cent of the signals in a given test sequence; then the average of these individual primaudibility ratios was computed. These two methods of calculating recognition differentials rarely gave results which differed by more than 1 decibel; such differences did not appear to vary in any systematic manner. Individual variability, however, is sometimes fairly large. Thus, it was noted in one of the tests that the entire transition curve for the least responsive observer was shifted by nearly 4 decibels with respect to the curves given by the others; when a retest was conducted on the following day, his transition curve was only about 1.5 decibels below the average curve.

The members of the test group agreed that the 1,000-millisecond pulse was not heard in a sustained fashion when the signal-to-noise ratio was insufficient to assure 100 per cent detection probability. Only short segments of tone, very much less than 1,000 milliseconds in length, could be sensed continuously during a single presentation, and apparently such segments were heard during periods when the level of the masking noise momentarily fell below its average value. A similar indication of the existence of sampling effect was obtained in a few qualitative tests with a dit-dit signal consisting of two 2-millisecond pulses separated by about 1 second. These repeated pulses were much easier to detect than single pulses of equal amplitude and duration because, when one of the pair was blanketed by a nearly simultaneous noise burst,[g] the other was usually clearly audible. The more audible of the two was just as often the first as the second member of the pair; in other words, there was no evidence that the ear integrated the energy of the two pulses.

8.3 BRITISH TESTS

The British have obtained useful information by means of the "echo injection" scheme. This consists of injecting a rectangular pulse at some point in the circuit of an echo-ranging receiver. If this receiver is mounted aboard ship, the type and level of masking noise are exactly what would be encountered in the field under the same conditions of operation. The signal may be injected either before or after the heterodyne stage and is presented to the ear at the usual output frequency, approximately centered with respect to the frequency limits of the admitted, masking noise band. An audio frequency of 1 kilocycle was used for the pulse in the tests described in the present section, and the spectrum of the heterodyned noise was very nearly flat.

In practice, echo-injection tests are conducted with the listening vessel either adrift or underway at various speeds; the received noise being picked up at representative bearings of the receiving hydrophone. In most cases, measurements are made of the amplitude of the primaudible injected pulse, but not of the level of the masking noise. However, if the listening conditions remain constant (pulse length, gain and filter settings, heterodyne frequency), such measurements, of the variation with speed of the primaudible amplitudes of injected signals, give operationally useful information in that they reveal how much stronger primaudible

[g] It is possible, of course, that the difference in audibility between two such pulses is related to the phase (see Figures 8 and 9 in Chapter 9) of the masking background as well as to its amplitude.

target echoes must be at one speed of the searching vessel than at another.

In contrast to this procedure, the data shown in Figure 7 were obtained[5] by measuring both the noise and the primaudible signal levels at the input to the headphones worn by the observers; consequently a critical band RD may be computed, and this has been entered as a lozenge in Figure 8. This value is in good agreement with the general trend of RD values shown in that figure. In fact, the same critical band RD was obtained for the 50-millisecond signal when the pulse-background mixture was passed through any of four filters whose audio admittance bands were 230 to 2,800 cycles (indicated in Figure 7), 500 to 1,400, 800 to 1,200, and 940 to 1,070 cycles respectively. In other words, the strength of the just detectable echo, for given conditions of self-noise, was not diminished by restricting the width of the masking noise band, because the components of the noise which masked the 1-kilocycle pulse were essentially those contained in the 50-cycle critical band centered at 1 kilocycle, and these noise components were transmitted equally well by all four of the experimental filters.

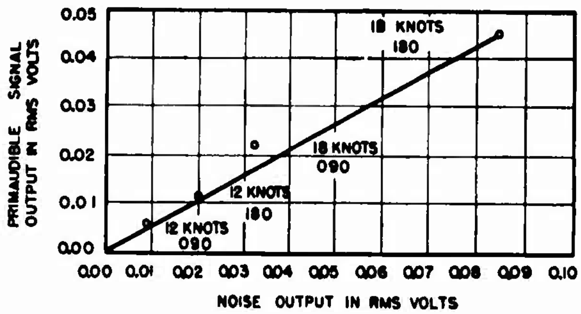


FIGURE 7. Variation of primaudible signal level with the level of self-noise, at indicated speeds and relative bearings, for a fixed gain setting. Heterodyned noise was presented in a band which extended from 230 to 2,800 cycles. The signal was a 50-millisecond rectangular pulse, heterodyned to 1 kilocycle. The slope of the line, or the signal-to-noise amplitude ratio, equals 0.530. Therefore the signal-to-noise intensity ratio equals $(0.530)^2$, or −5.5 decibels [20 log 0.530]. The equivalent ratio for a 50-cycle noise band is greater by 17.1 decibels [10 log (2,570/50)]. Hence the critical band recognition differential equals 11.6 decibels.

The points in Figure 7 represent averages obtained from a large number of observations. This procedure was found necessary because of the fairly high degree of variability of received self-noise even for apparently identical operating conditions.

It will be seen that the amount of masking is directly proportional to the level of the masking noise, and the masking vanishes when the noise vanishes. Furthermore, the linear relation shown in this figure indicates that the masking efficiency of the noise was identical for each of the four combinations of speed and bearing specified; even when the noise has a prominent amplitude modulation (as in the case of the 180-degree orientations), the only feature of the masking background which is statistically significant is the mean value of the rms noise level.

It may be inferred, therefore, that target noise, which often has a large degree of amplitude modulation when the target is a surface vessel, will not have more masking efficiency than, for example, deep-sea ambient; field experience[3] has shown that this is true, even for pulses as short as 10 milliseconds. It was not true, however, for nonauditory methods of detection in which intensity is the major cue and in which masking of pulses becomes more effective when the noise has a high peak factor.

It should be added that primaudible signal-to-noise ratios were determined, in these tests, by the method of minimal increments. Such primaudibility values usually exceed the 50 per cent recognition differentials obtained from transition curves by about 1 to 2 decibels (see pp. 188, 231), and it may be partly due to this fact that the lozenge at 50 milliseconds tends to lie a little above the line drawn through the other filled-in symbols. It is also worth noting that the report describing the tests under discussion states that the gain was set for a comfortable listening level and that the primaudibility ratio changes somewhat when the listening level is varied over a wide range. It is not known whether this variation was produced by the alteration of critical band width with sensation level, by the greater importance of environmental noise at low listening levels, by distortions due to overloading, or by some other factor.

It was also found that hydrophone effect (listening for propeller sounds received by the echo ranging transducer) obtained on the screws of a vessel running a course parallel to that on which the tests were performed began to suffer when the listening band width was reduced much below 800 cycles. In other words, when the band width is insufficient to admit the fundamental and prominent harmonics of the modulation wave form, the character of the signal modulation is poorly defined and may become unrecognizable.

Band widths greater than 800 cycles gave no advantage in listening either to screw sounds or to pulses. In fact, admission of frequencies above 1,400 cycles annoyed the listeners to some extent because the sound became harsher. As in most supersonic listening, the presented spectrum was flat and the sensation levels of the high-frequency components were relatively high; hence, the large degree of annoyance.

8.4 UCDWR TESTS ON SINGLE PULSES

Some of the descriptive material in the present section will be found in reference 6. The remainder of the data collected here have been privately communicated by UCDWR in advance of publication, which is expected in the near future; for more complete information, the forthcoming report on this subject, to be issued by UCDWR, should be consulted.

8.4.1 General Procedure

Recognition tests were conducted in the presence of thermal or recorded self-noise with isolated pulses and with strings of pulses repeated at various rates. The work with isolated pulses is discussed first. The general test procedure used in either case was the following. Mixtures of signal and background were presented to groups of 3 to 5 observers by means of high-quality headphones, in a fairly quiet room, and at a comfortable listening level. Gain applied to the signal-background mixture could be adjusted for maximum comfort by the individual observers. The signal-to-noise ratios covered the range from nearly zero to nearly 100 per cent recognition probability in steps of 2 decibels. In any test sequence, gain in the background channel was held fixed and the various signal-to-noise ratios were equally and randomly distributed among the successive listening intervals. The observers expressed their judgments by means of a well-adjusted hand key which could be switched from a neutral position to either of two others corresponding to audible and inaudible respectively. Transition curves were plotted for all observers whose responses showed no significant number of commissive errors for the blank presentations, and the 50 per cent values so obtained were used to represent primaudible signal-to-noise ratios. Rest periods were taken between test sequences in order to minimize fatigue.

None but CW signals were studied in these laboratory tests. Details of pulse length and band width are given below in connection with specific test results; so also are statements regarding deviations from the general procedure just described, for example, self-administered tests.

In general, values of n for transition curves obtained in these tests fell between 2 and 3 and did not appear to depend significantly on background type nor on whether the signal was an injected pulse or a recorded echo. The method of minimal increments gave recognition differentials larger by 1 to 2 decibels than obtained with the standard test procedure. Comparisons made in a few cases between transition curves and recognition differentials, obtained for given signal-background mixtures by means of either high-quality loudspeakers or headphones, showed no significant advantage for either method of presentation under the rather favorable laboratory conditions used.

8.4.2 Rectangular Pulses

A heterodyne frequency of 800 cycles is standard in current American echo-ranging gear; hence, all recognition differentials obtained with pulses of or near this frequency have been collected in Figure 8 and indicated by means of filled-in symbols. The UCDWR tests add several points to this graph, three of them (for rectangular pulses 1, 149, and 4,000 milliseconds in duration) being shown as solid circles. A typical rectangular pulse used is shown in Figure 9. The masking background in these cases was thermal noise whose

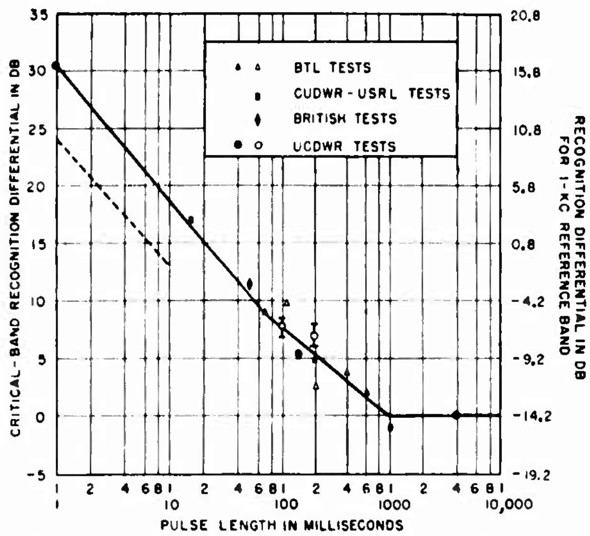


FIGURE 8. Aural detection of rectangular 800-cycle pulses masked by wide-band thermal noise.

spectrum was nearly flat and which extended from 0.1 to 10 kilocycles for the tests with the 1- and 149-millisecond pulses. In the case of the 4-second pulse, the masking noise was admitted by one of two band-pass filters, transmitting either one octave (565 to 1,130 cycles per second) or two octaves (400 to 1,600 cycles per second); identical critical-band recognition differentials were obtained with both of these filters. A critical band width of 38 cycles has been used for all calculations involving 800-cycle pulses. It will be noted that the essential spectrum of a 1-millisecond pulse extends over an interval of 2,000 cycles (2/0.001). Consequently, a wide-band background is required to mask all the components in the pulse spectrum, and conversely, a wide-band system is required to pass a 1-millisecond pulse without substantial distortion.

In tests described in Section 8.5, rectangular 800-cycle CW pulses with durations between 1 and 10 milliseconds were repeated at various rates and presented against thermal noise backgrounds extending from 0.1 to 10 kilocycles. When a pulse of given duration was repeated at a regular rate (5.6 times per second), the critical-band recognition differentials for the strings fell along the dotted line shown in Figure 8. This line is included in the figure for comparison with the slope of the line drawn through the filled-in symbols, since the latter are somewhat sparse in the region of short pulse lengths. It will be observed that these lines have nearly identical slopes and that the recognition differentials for repeated pulses are about 6 decibels smaller than for single pulses.

8.4.3 Rounded Pulses

The envelopes of these pulses (with durations of 100 and 200 milliseconds, respectively) are shown in Figure 9. In the absence of masking, these pulses had a slightly less "fuzzy" sound than did the rectangular pulses; their objective spectra are narrower or, what amounts to essentially the same thing, they

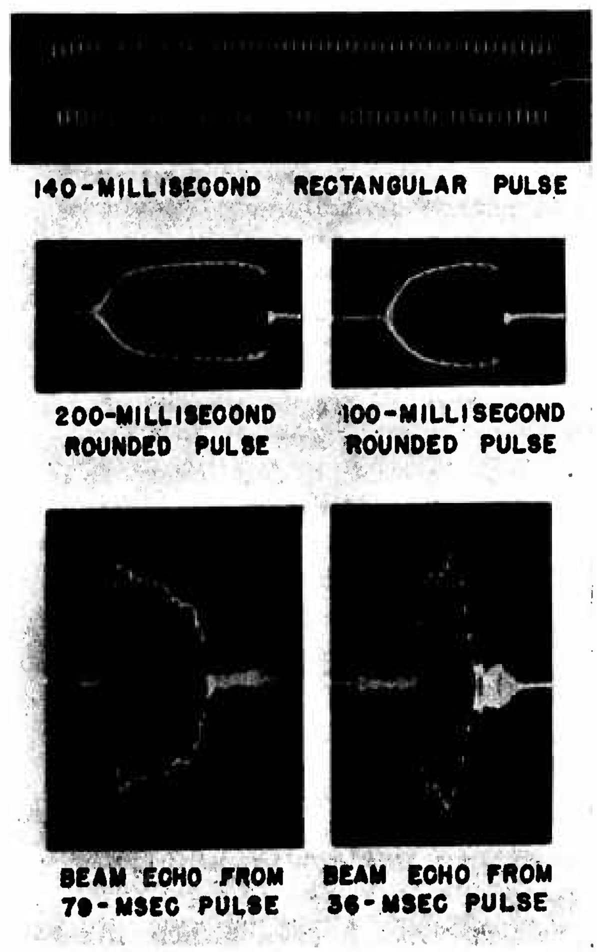


FIGURE 9. CFO traces of rectangular and rounded CW pulses, and recorded echoes.

generate less click transient within the ear.

The masking background used in these tests was thermal noise with a flat spectrum, admitted by a band-pass filter with cutoffs at 400 and 1,600 cycles per second. Critical-band recognition differentials for the rounded 100-millisecond and 200-millisecond pulses are shown as open circles in Figure 8. The vertical lines drawn through these circles represent the difference in performance between the best and poorest observers. These indications of the degree of scatter (which are fairly typical for rectangular as well as rounded pulses) have been included for comparison with the scatter in performance usually observed with recorded echoes and described in Section 8.4.6.

It will be noted from Figure 8 that the rounded 200-millisecond pulse was about 2 decibels less audible than a rectangular pulse of the same duration. This is in agreement with the estimated loss due to rounding which has been discussed in connection with Table 2 and implies that the sharp onset and termination of a rectangular pulse (the click transient) provides a useful cue in distinguishing signals from false indications. It is not known to what degree the pulses emitted by practical sonar projectors are rounded, but it hardly seems likely that the majority of echoes returned by underwater objects will have ideal rectangular envelopes even when the outgoing ping does have such a characteristic. It is consequently helpful to have the results shown in Figure 8, in order to form some conception of the manner and degree by which recognition of injected rectangular pulses differs from that of echoes encountered under service conditions.

8.4.4 Recognition Differentials for CW Pulses

The available data on the relative audibilities of rectangular 800-cycle pulses of various durations are all plotted in Figure 8. Among the filled-in symbols, the triangles refer to 740-cycle pulses and the lozenge to a 1-kilocycle pulse; the others, to 800-cycle pulses. As mentioned in Sections 8.2 and 8.5, variations in observed signal-to-noise ratios required for primaudibility are almost negligible for pulses of these frequencies. Hence, the results are grouped in Figure 8 and collectively designated as applying to 800-cycle pulses.

Although the plotted data are drawn from many sources, and were obtained under somewhat different conditions of test, they appear to define a single function. The filled-in symbols (all points other than those obtained with rounded or modulated pulses) have been fitted provisionally by means of three straight-line segments, by analogy with Figures 3, 4, and 5, although possibly a curve, or curves, should be used. In any case, the data seem to require a function which is concave upward and which has a larger (negative) slope in the region of shorter pulse lengths. Independent evidence, already discussed in Sections 2.2.2 and 8.1.2, also implies that the auditory function relating *RD* and τ is not a single straight line in the region of pulse lengths below about 600 milliseconds.

The results presented in Figure 8 may be summarized as follows. For signals between 1 and 10 seconds in length, and presumably also for greater signal lengths, the critical band *RD* is zero; thus, a long sustained tone can be recognized 50 per cent of the time when its power is just equal to the noise power in the corresponding critical band. For a 1-kilocycle reference band the *RD* is -14.2 decibels for an 800-cycle signal (10 log 38/1,000). For signals less than 1 second in length the critical-band *RD* increases (becomes less favorable) with decreasing signal length, with an overall change of 30 decibels as the signal length is diminished to 1 millisecond. This change is primarily due, of course, to the finite build-up time of the ear. However the data seem definitely to suggest a change in slope in the neighborhood of 50 to 100 milliseconds, with slopes of 12 decibels and 8 decibels per tenfold increase of pulse length for the shorter and longer pulses. The data are not sufficiently accurate, however, to establish the reality of this break in slope.

These results are somewhat different from those obtained in the study of the loudness of pulses. A comparison of Figure 8 with Figures 3, 4, and 5 shows in fact two major differences. In the first place, the intensity of the primaudible pulse, in the presence of noise background, increases as the pulse length decreases below 1 second, while for constant loudness, in the absence of masking, the intensity remains constant, down to 250 milliseconds. Thus, if the primaudible 250-millisecond pulse is compared with the primaudible 1-second pulse, in the absence of masking, the shorter pulse will sound louder than the longer one. In the second place, the slope of the *RD* curve shown in Figure 8 is less than the slopes shown in Figures 3, 4, and 5. Thus, the intensity of the just audible pulse, in the presence of noise background, does not increase as rapidly with diminishing pulse length as does the intensity of the pulse of constant loudness. Or, in other words, the loudness of the primaudible pulse, heard in the absence of the masking background, apparently decreases with decreasing pulse length. Since loudness and masking have usually been found to be directly related, these results are somewhat surprising and a provisional explanation of the discrepancy is given below in terms of the sampling effect.

The increase of *RD* found for pulse lengths between 0.25 and 1.0 second may be attributed at least in part to the variability of the masking noise background. The longer the pulse, the greater the probability that the noise level will fall sufficiently far below the average to make possible the recognition of the pulse. This phenomenon, which has already been referred to as "sampling effect," is also suggested by the results reported in Section 8.2 in which it was noted by the observers that the primaudible 1-second pulse could not be heard for its entire duration but only during relatively low intervals in the background noise.

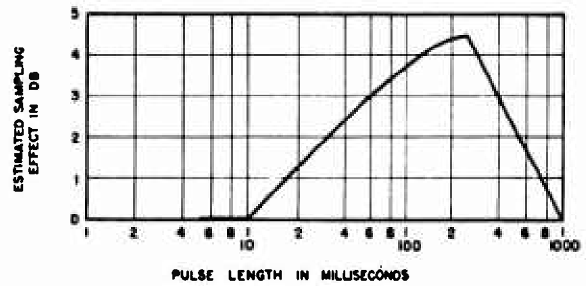


FIGURE 10. Deterioration of audibility, partly due to the sampling effect.

A rough estimate of the quantity of interest is shown in Figure 10, obtained by plotting the difference between the recognition differentials found from Figure 8 and the intensity increments required for constant loudness shown in Figure 4. The latter applies to 1-kilocycle pulses, but the properties of 1-kilocycle pulses are probably very similar to those of 800-cycle pulses. The data corresponding to a sensation level of 20 phons have been taken from Figure 4. If the 60-phon data were used, the curve would be the same for pulse lengths between 0.25 and 1.0 second but would come to zero more rapidly for the shorter pulses. It seems clear that the sampling effect should explain some of the difference shown in Figure 10, although the actual importance of this effect is, of course, uncertain.

The disagreement between the slopes for the loudness and the masking curves has no immediately obvious explanation. It has already been noted in Section 8.1.2 that the results shown in Figures 3, 4, and 5 are not too reliable, since the basic data on which they are based show an appreciable scatter. Thus it is not inconceivable that an intensity change of 10 decibels per tenfold change in signal length may possibly characterize both the pulse of constant loudness and the pulse which can just be heard above the background noise.

One additional discrepancy may be noted between the masking data and the constant loudness curves shown in Figures 3, 4, and 5. The data in Figure 8 were obtained with nine different noise bands, ranging in width from 0.3 to nearly 10 kilocycles. In all cases the loudness of the noise background was adjusted to a value of about 70 phons. Thus, the noise energy contained in a critical band centered at 800 cycles varied from about 20 decibels above the audibility threshold to about 65 decibels above this threshold. Despite this large difference in critical-band listening level, no systematic change of *RD* with band width was noted. This is in marked contrast with the data shown for 1-kilocycle pulses in Figure 4, where a considerable dependence on loudness level was found. The data are, again, not sufficiently precise to yield reliable results; nevertheless, the practical conclusion seems indicated that the recognition differential is independent of listening level for practical values of that level.

The preceding discussion may be summarized as follows. The masking data shown in Figure 8 seem to show significant differences from the data for pulses of constant loudness. Some of these differences may perhaps be explained in terms of the so-called "sampling effect." Others are less readily explained and may or may not be real. It is entirely possible that the response of the ear may be quite different for very short pulses in the presence of masking noise, and for the same pulses in the absence of background.

An effect already briefly discussed in Section 8.1.2 may be considered next, namely, the broadening of the pulse spectrum beyond the limits of a single critical band. For an 800-cycle tone, the critical band is 38 cycles wide, and for a 50-millisecond pulse the essential spectrum width is 2/0.05, or 40 cycles. Thus, for pulses of given intensity which are less than 50 milliseconds in length, the signal energy in each critical band decreases with decreasing pulse length. By way of illustration, consider a 4-millisecond pulse whose essential spectrum has a width of approximately 500 cycles (2/0.004), and which therefore extends from a frequency of about 550 cycles to one of about 1,050 cycles (800 ± 250 cycles). Such a spectrum stimulates at least ten contiguous critical bands, since no critical band in this frequency region is more than 50 cycles wide. It follows that the computed signal-to-noise ratio in each critical band is on the average about one-tenth of that shown in the figure, that is, 10 decibels lower. Since the distribution of pulse energy over the essential width of the spectrum is not uniform, the signal-to-noise ratio for critical bands near the center of the pulse spectrum will exceed this average value. From equation (3) of Chapter 7 it may be shown that the sound power per cycle at the midfrequency is twice as great as the average sound power per cycle in the "essential spectrum" of width $2/\tau$ cycles. For present purposes, however, this complicating factor may be neglected, and the average sound power is assumed to be uniformly distributed over the essential spectrum. This assumption appears reasonable when there is some cooperation between the different critical bands stimulated by a short pulse, and will in any case not be in error by more than 3 decibels. On this basis, then, the corrected value of the critical band *RD* may be considered as applying to any of the critical bands stimulated by the pulse spectrum, since the masking noise was essentially flat in all the pulse tests; in other words, signal and noise spectra had approximately the same shape.

Thus, an estimate may be made of primaudible signal-to-noise ratios for individual critical bands when the pulse spectrum occupies several critical bands and parallels the noise spectrum. It was shown above that this correction is of the order of 10 decibels for a 4-millisecond pulse. Similar corrections have been applied in the case of pulses between 1 and 40 milliseconds in length. The corrected results

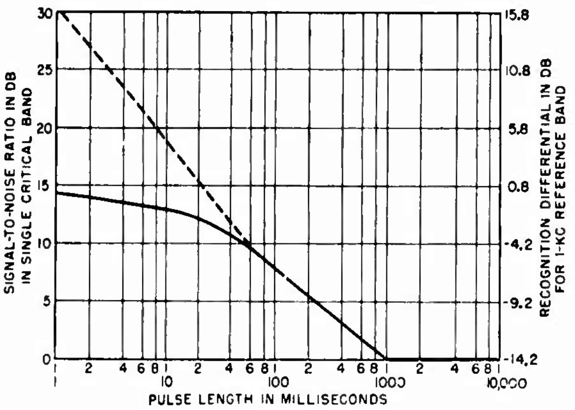


FIGURE 11. Critical band recognition differentials for 800-cycle pulses masked by wide-band thermal noise.

are shown in Figure 11, where the dotted line is transcribed from Figure 8 and represents the critical-band recognition differentials computed without regard to the width of the pulse spectrum. Clearly, the signal-to-noise ratio in a single critical band tends to remain nearly constant for pulses which are so short that their spectra stimulate more than a single critical band, but this approximate constancy per critical band is equivalent to an actual loss in performance because of the progressive broadening of the pulse spectra with diminishing duration. This application of the critical-band picture must be regarded as tentative at the present time, especially in view of its approximate nature. It is shown in Sections 8.5.3 and 9.2.2, however, that this extension of the critical-band concept agrees with other experimental data.

For practical purposes, it is most convenient to present the data in the form used in Figure 8 and to express the recognition differentials in terms of the 1-kilocycle presentation band on which the right-hand ordinates are based. This procedure is followed in the remainder of the discussion, because the effects of changing such factors as heterodyne frequency or system band width are most easily assessed by referring all primaudible signal-to-noise ratios to the same standard. Thus, the use of the 1-kilocycle reference band in plotting points for the FM pulse data shown in Figure 8 shows much more clearly than could otherwise be done whether this procedure might be expected to help in the field.

8.4.5 Effect of Heterodyne Frequency

Inasmuch as the heterodyne frequency used in practice can be selected arbitrarily within rather wide limits, it is desirable to evaluate the effect which changes in this frequency can produce on overall performance. Changes in audio frequency might be expected to produce important changes in the masking of signals and in the precision with which doppler shifts can be recognized and identified. In addition, operator fatigue may be expected to depend on the heterodyne frequency used. The data presented in this section are too preliminary to give conclusive results but are believed to shed useful light on the effects to be expected.

Recognition tests were conducted with rectangular pulses of three frequencies (0.4, 2, and 6 kilocycles) in the presence of a nearly flat thermal noise band extending from 0.1 to 10 kilocycles. The pulse lengths used in this work varied from 1 to 147 milliseconds. The test procedure followed was, in general, the same as that described in Section 8.4.1. The only significant deviation was that all the masking data for pulses of 400 cycles, and some of the data for 6-kilocycle pulses were obtained in a self-administered test by an experienced observer. All the masking data at 2 kilocycles and some of the data for 6 kilocycles were determined in the usual manner with a number of listeners. General observations at UCDWR indicate that the results obtained in the self-administered test are not likely to differ by more than 1 or 2 decibels from the recognition differentials found in the more usual way. The same conclusion is evidenced by a comparison of the two types of results obtained at 6 kilocycles and plotted in Figure 13.

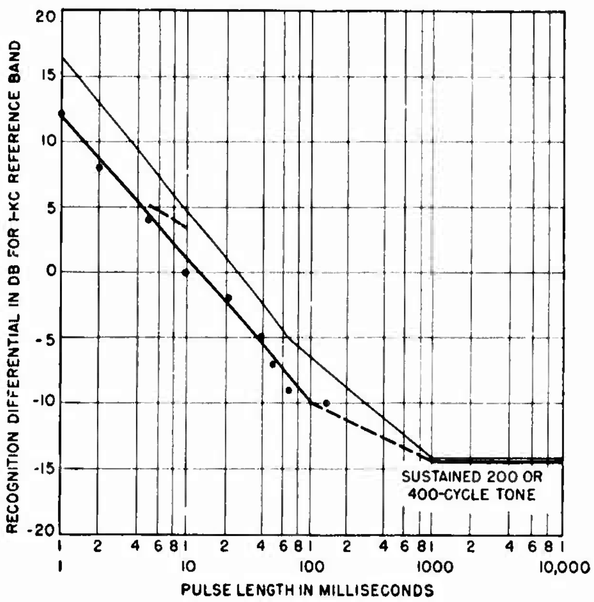


FIGURE 12. Audibility of 400-cycle pulses masked by thermal noise. The thick line, a visual best fit of the observed points, represents recognition differentials for 400-cycle pulses in a self-administered test. The thin line, taken from Figure 8, gives recognition differentials for 800-cycle pulses. The upper dotted line represents recognition differentials for 200-cycle pulses, repeated 5.6 times per second.

In a series of self-administered tests, the recognition levels were found for rectangular 400-cycle pulses of different lengths. Masking was produced by thermal noise whose spectrum extended from 0.1 to 10 kilocycles. The recognition differentials found, referred to a reference band 1 kilocycle in width, are plotted in Figure 12. The upper solid line in this figure is the mean curve taken from Figure 8, showing the corresponding recognition differentials for rectangular pulses of 800-cycle sound. Since no data were obtained at 400 cycles for pulse lengths between 150 milliseconds and 1 second in length, the dotted line in this region is intended only to carry the eye across this gap. The true position of the curve in this region is, of course, unknown.

Figure 12 apparently indicates that 400-cycle pulses become primaudible when they are 3 to 4 decibels weaker, relative to the adjacent noise background, than 800-cycle pulses of the same length. Although the slope shown for the 400-

cycle curve is probably reliable, there are two reasons for questioning the absolute advantage indicated for pulses of this lower frequency. In the first place, the recognition differentials for sustained tones of 400 and 800 cycles differ by only a fraction of a decibel. In the second place, several lines of evidence indicate that for very short pulses there is little difference in the detectability of signals of 400 and 800 cycles. For example, the tests discussed in Section 8.2 showed that for pulses of 16 milliseconds duration, the recognition differential was the same for pulses of 400, 800, and 1,600 cycles. Similarly, in the tests on repeated pulses described in Section 8.5, no change of recognition differential was observed for a change of frequency between 0.5 and 6 kilocycles and for pulse durations of less than 10 milliseconds.

This similarity between the recognition differentials of very short pulses of different frequency is generally consistent with the fact that such short pulses sound much the same regardless of the frequency. The essential spectrum of a pulse 1 millisecond in length is at least 2,000 cycles wide. Thus, it is to be expected that changes of center frequency as great as 1,000 cycles would have little effect on *RD.* It is evident physically that when a pulse is roughly 1 cycle long, the frequency of the pulse ceases to have very much significance.

These expectations are confirmed by experiments on the observed pitch of short pulses. These observations were made in connection with the tests on repeated pulses, described in Section 8.5. Tonal pulses of various lengths were compared with noise pulses of the same lengths. The noise pulses were obtained from recorded water noise passed through a filter centered at 1 kilocycle. The band width of this filter was adjusted to admit a number of cycles equal to the reciprocal of the pulse length in seconds. The various tone and noise pulses were presented to several listeners both with and without wide-band masking background. Even when no background was present the observers found it impossible to distinguish between a pulse of pure tone and a pulse of water noise when these had a duration of less than 1 millisecond. Tonal pulses 5 to 10 milliseconds long gave a slight but rather indefinite semblance of pitch. Variation of the frequency of the tonal pulse from 700 to 1,000 cycles did not change these results. Similar practical observations made during field tests[3] gave essentially the same results.

It may be inferred, therefore, that the points plotted in Figure 12 for very short pulses should agree with the curve for 800 cycles. Thus, the evidence strongly indicates that some systematic error affects all the points plotted in this figure for 400-cycle pulses. Possibly inadequate weight was given in this work to variations with frequency in the spectrum level of the masking noise (see Section 8.5.2). A part of the discrepancy may have arisen from the fact that the 400-cycle tests were self-administered.

The upper dotted line shown in Figure 12 is taken from Figure 19 and shows recognition differentials obtained with a 200-cycle sequence repeated 5.6 times per second. This line is inserted here for subsequent reference in Section 8.5.3.

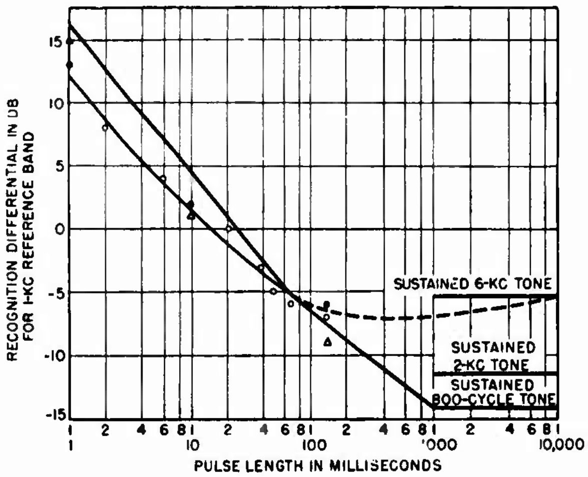


FIGURE 13. Audibility of 2- and 6-kilocycle pulses masked by thermal noise. The curve, a visual best fit of the circles, represents recognition differentials for 6-kilocycle pulses; the other line, taken from Figure 8, gives recognition differentials for 800-cycle pulses. The open circles are observations made in a self-administered test, while the filled-in circles were obtained in a test administered to three observers. The triangles indicate recognition differentials for 2-kilocycle pulses in a test given to three observers.

Recognition differentials obtained with rectangular pulses of 2 and 6 kilocycles are plotted in Figure 13. The masking background was again wide-band thermal noise, and the recognition differentials are referred to a 1-kilocycle reference band as in Figure 8. The open circles represent the data obtained in a self-administered test at 6 kilocycles, while the filled-in circles represent the corresponding data obtained with three observers in the usual manner. Evidently any systematic difference between these two sets of points is less than 2 decibels. One of the lines in Figure 13 again represents the recognition differentials found at 800 cycles, taken from Figure 8. The curve has been fitted to the 6-kilocycle points, and, as before, the dotted line has been sketched in to carry the eye across the region from 150 milliseconds to a sustained tone of 1 second or more. No curve has been drawn for pulses of 2 kilocycles because of the paucity of the data.

In Figure 13, as in Figure 12, the points for the shorter pulses lie systematically some 3 to 4 decibels lower than the curve for 800 cycles. Since the data for 400 cycles, 2 kilocycles, and 6 kilocycles were all obtained in the same test series, it seems likely that the recognition differentials for all three frequencies are affected by the same systematic error. It has already been pointed out above that the recognition differentials for pulses about 1 millisecond in length should be independent of frequency in the range between 0.4 and 6 kilocycles. If all the points shown in Figures 12 and 13 are shifted upward by 4 decibels, they will be in agreement with the 800-cycle data for the shortest pulses. This does not appear to be an excessively large shift in view of the many possibilities of systematic error in psychoacoustic work.

When the recognition differentials for pulses of 0.4, 2, and 6 kilocycles are increased in this manner, all the available observations seem to be in fairly good agreement. Figure 14 has been drawn to give estimated recognition differentials at 0.4, 0.8, 2, and 6 kilocycles on this basis. Although these curves can scarcely be regarded as experimentally established, they are probably the best that can be drawn on the basis of present evidence. The curves shown for the data at 2 and 6 kilocycles have been drawn smoothly. Abrupt changes in slope might be expected at 15 and 10 milliseconds where the essential pulse spectra become equal in width to the critical band widths at 2 and 6 kilocycles, respectively. However, the evidence is insufficient to indicate such discontinuities in slope.

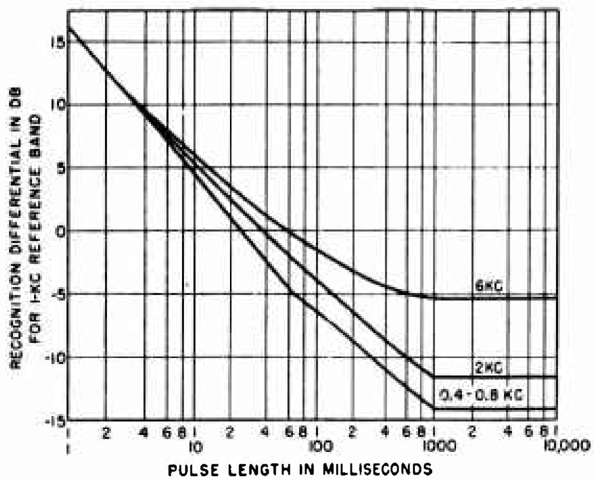


FIGURE 14. Estimated recognition differentials for pulses of different frequencies, masked by wide-band thermal noise.

It may be observed that the estimated curve for pulses of 6 kilocycles has a more gradual slope between 10 and 100 milliseconds than the curves for the other frequencies. This is in part due, of course, to the fact that for long pulses, the recognizable 6-kilocycle signal must be higher relative to the background than for lower-frequency signals of equal duration. It is possible that changes in the sampling effect with changing frequency may also be important in explaining the difference between the curves for different frequencies. Owing to the increased width of the critical band at 6 kilocycles, the total noise level in this band will vary more rapidly (see Figure 57 in Chapter 4). The effect of this variability on signal recognition will depend on the details of the hearing mechanism. If, for example, the perceived sound represents a time average of the sound received in a critical band, the variability of the noise level will be decreased as the critical band width increases. On this basis the sampling effect should decrease in importance with increasing critical band width. Definite conclusions on this subject, however, must await more extensive investigations of the performance of the ear.

The preliminary curves shown in Figure 14 appear to justify, from the standpoint of aural masking, the general choice of heterodyne frequency in current American and British sonar gear (0.8 and 1 kilocycle, respectively). Probably any frequency between 400 cycles and 1 kilocycle would provide the same recognition differential for echoes in the presence of a noise background. The overall justification of the heterodyne frequency used in practice must, however, take many other factors into account. Among these other factors are the recognition differential for echoes in the presence of reverberation, discussed in Chapters 9 and 10, and the ability to distinguish small doppler shifts of the echo relative to the reverberation.

8.1.6 Recorded Echoes

The echo recognition tests described in the previous sections used rectangular pulses artificially produced. Since there are many differences between such ideal pulses and the actual echoes obtained from submarines, a limited series of observations was also carried out with recorded echoes. The echoes used were obtained by echo ranging at sea on a submerged S-class submarine several hundred yards away. Film recordings were made of the received signal, heterodyned to 800 cycles, and relatively free of background noise and reverberation. For the tests described here, the submarine was kept as close to beam aspect as possible. Under these conditions[h] the returning echo is usually similar to the outgoing pulse, though rarely an exact reproduction. Thus, the results might be expected to be somewhat similar to those obtained with rectangular pulses. No tests have been made on the noise masking of the "smear" type of echoes usually obtained at bow, stern or quarter aspect.

During the masking tests these recorded beam-aspect echoes were mixed electrically with thermal noise recorded on a separate film loop. The use of the recorded noise background gave a more nearly uniform noise level than can generally be obtained with a thermal noise generator. The mixture of recorded echo and thermal noise was passed through either a one-octave (565 to 1,130 cycles) or a two-octave (400 to 1,600 cycles) band-pass filter. The width of these filters seems to have had no significant effect on performance in the masking tests.

Test sequences were conducted using one individual echo recording at each time. In each test the noise and echo film loops were run continuously; the echo was thus injected into the noise at regular intervals of 3 to 5 seconds, depending on the length of the film loop. Since noise and echo loops were of different lengths, an echo was rarely repeated at precisely the same time in the noise loop.

Noise level was measured by means of a thermocouple or a copper oxide type meter. In all cases, signal level was defined in terms of the peak amplitude of the echo. This was done by using an oscilloscope as a visual indicator for adjusting the amplitude of a sustained tone to the peak amplitude of the echo. The level of the sustained tone was then determined by means of a meter. In a few measurements of longer echoes made by means of integrating meters, it has been found that the average echo level is between 1.5 and 2 decibels below its peak level.

In making each masking test, the signal-to-noise ratio was varied in a random fashion by means of an attenuator in the signal channel. Blank intervals were occasionally provided for evaluating the number of commissive errors. The scoring technique used was similar to that described in Section 8.4.1. Transition curves were plotted for groups of five observers, and the recognition differentials are shown in Figure 15. These are composite points, showing the performance of the group rather than that of the individual observers. As pointed out in Section 9.2, the difference between performance of the best and poorest members of the group may be as large as 7 decibels when real echoes are used. The circles in Figure 15 represent the average of all determinations made with individual echoes; since these are averages of decibels, they correspond to geometric rather than arithmetic means.

In general, Figure 15, shows the same trend as shown in Figure 8 for rectangular pulses; thus there is a loss in audibility of somewhat

[h] See Division 6, Volume 8, Chapter 23.

less than 10 decibels for a tenfold reduction in pulse length. It will be noted that there are apparently significant differences between the recognition differentials for artificial pulses and for echoes. Such differences may be expected as a result of the time-amplitude pattern of the observed echoes.

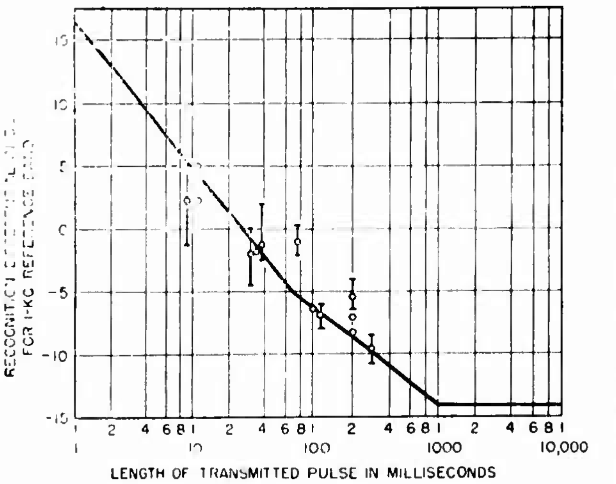


FIGURE 15. Aural detection of beam echoes masked by thermal noise; echoes were recorded after heterodyning to 800 cycles. Circles represent mean recognition differentials for specific echoes, and vertical lines indicate average scatter among these mean recognition differentials. The curve, taken from Figure 8, gives recognition differentials for 800-cycle pulses.

In the first place, the average intensity of a real echo several hundred milliseconds in length is, in general, several decibels less than the peak intensity used in computing the *RD*. This would tend to give larger (less favorable) recognition differentials than for rectangular pulses whose peak and average intensities are equal. For shorter pulses, the difference between peak and average intensity is apparently less, and this effect is presumably of diminished importance. In the second place, real echoes tend to be somewhat prolonged compared with the outgoing pulse. This effect is less important for beam echoes than for those at other aspects. Even at beam aspect, however, the returned echo may be 5 milliseconds or so longer than the outgoing pulse. For very short pulses this obviously leads to a considerable prolongation of the echo. A longer echo is easier to hear than a shorter one. Since the echo length used in Figure 15 is actually the length of the emitted pulse, rather than the length of the observed echo, echo prolongation will therefore tend to give smaller (more favorable) recognition differentials than the rectangular pulses of fixed length.

Difference of tonality between real and artificial echoes may also be responsible for differences in *RD*. The tonality of real echoes tends to be inferior to that of pulses. This effect is in part the result of amplitude variations of the echo envelope but may also be due in part to abrupt changes of phase occurring during the echo. In general, echoes heterodyned to 800 cycles had definite tonality in these tests only when they were more than 10 milliseconds in length. A shorter echo sounded like a crack or pop. Even fairly long echoes tended to sound "noiselike" when their envelopes were very uneven. Another difficulty encountered in the recognition of real echoes is the absence of the click associated with rectangular or nearly rectangular pulses. This type of cue seems to be more helpful in the case of the longer pulses, since the shorter have little tonality and are mostly click in any event.

The plotted points shown in Figure 15 are consistent with these theoretical expectations. For example, two points are shown for 9-millisecond pulses. One of these was actually 18 milliseconds long, as determined by measurements made on the film. The mean *RD* for this 18-millisecond echo, shown by the lower of the two points plotted at 9 milliseconds, is fairly close to the *RD* indicated for a rectangular 18-millisecond pulse. The other echo was observed to be about 9 milliseconds long, and its observed *RD* is much closer to that for a rectangular 9-millisecond pulse. Similarly, the general trend in Figure 15 shows a greater loss in audibility relative to the rectangular pulse data for the longer echoes than for the shorter. This would be expected from the increasing difference between peak and average intensity as pulse length increases, and also from the increasing loss of tonality relative to rectangular pulses. Figure 8 shows a similarly more gradual slope for rounded pulses as compared with rectangular ones. In view of the wide scatter of the data, however, this agreement with expectation must be regarded as suggestive rather than conclusive.

The vertical lines through the circles in Figure 15 show the average degree of variability in group recognition differentials which were obtained in successive tests with particular echoes. As indicated above, the total difference between the best and the worst performances by individual observers during an entire test sequence is even greater. Although these studies were of a preliminary nature only, ten or more successive tests were made with some of the echoes (such as the 18-millisecond echo discussed previously). No progressive changes in performance were found to occur from test to test. It may be, however, that further experience would have reduced the variability to some extent, since constant practice seems required for the maintenance of peak performance in listening tests with real echoes (see Section 9.2.2). This would be anticipated from the fact that real echoes are objectively more complex than rectangular CW pulses; there are more cues which a skilled observer can rely upon and, hence, more variability among observers. Compare, for example, the variability indicated in Figure 15 with that shown in Figure 8.

The shapes of transition curves obtained in this study with real echoes were for the most part about the same as observed in tests with rectangular CW pulses. The only major differences observed were for the recognition of the 9-millisecond echoes, in which case the slope of the transition curve was very much more gradual than for CW pulses of equal length. This slow rate of improvement with increasing signal level is probably related to the factors producing general variability in echo recognition differentials; cues are ill-defined, hence increases of intensity are not as helpful as with rectangular pulses.

In conclusion, it should be emphasized that echoes received from bow and stern aspects of a target would in general be even more smeared than the beam-aspect echoes used in the tests under discussion. In addition, even beam echoes would be expected to give different results when multiple transition paths can substantially prolong the observed echo. To obtain more general information on echo masking, studies are required with a wider variety of echo samples. In addition, it may be useful to obtain fundamental information on the effects produced by arbitrary variations in the envelope and phase of an artificial pulse.

8.4.7 Effect of System Distortion

Receiver output O may be related to input I in a very large number of ways. When O is directly proportional to I, the receiver is linear; in all other cases the receiver is nonlinear. Although linearity is frequently an ideal with electroacoustical equipment, it is rarely achieved over more than a limited range. Since various kinds of nonlinearity may occur in practice, it is important to inquire to what extent nonlinearity may affect the recognition of target echoes. Several tests along this line have been carried out by UCDWR.

Two important types of nonlinearity found in practice are represented in Figure 16. In

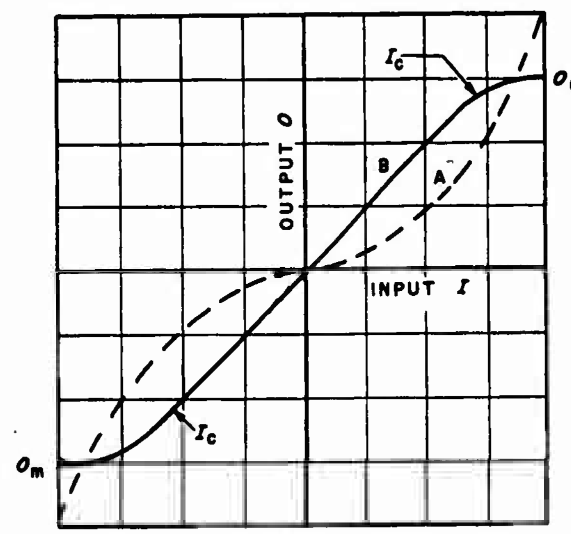


FIGURE 16. Two types of nonlinearity.

curve A, a change in I, when I is small, produces very little change in O; but with increasing I, the output becomes more and more sensitive to small changes in the input. This situation arises when I is proportional to O^2 as, for example, in doppler-doublers. In curve B, on the other hand, I and O are proportional if the input does not exceed the critical value I_c shown on the curves. Within this range of input the receiver is practically linear. When

I is increased above I_c, the output curve flattens off and approaches asymptotically a maximum output O_m. In this nonlinear region the receiver is said to be overloaded. Since the output will not exceed O_m, no matter how great the input, the receiver is said to have a limiting action. Nonlinearity of the type shown in curve A is important in square law amplifiers. Nonlinearity of the type shown in curve B is encountered among practical types of gear which are designed to be linear over a specified operating range. The UCDWR masking tests have been carried out only for receivers whose response is similar to that shown in curve B.

When the input in an echo-ranging receiver increases beyond the critical input I_c and the receiver overloads, several effects may be produced in the observed masking of echoes. This problem is particularly important when short tonal signals are to be detected in the presence of noise, since then the amplitude of the just audible signal often exceeds that of the background. This difficulty may be overcome by diminishing the total gain, so that the system operates along the quasi-linear portion of the curve below I_c on the solid line in Figure 16. When the signal amplitude is equal to or less than that of the admitted noise band, as is usually the case for long tonal signals, the difficulties produced by overloading and limiting may be minimized by diminishing system band width, and thereby the total amplitude of the signal-background mixture. When the receiver has a small dynamic range it is often suggested that the operator set the gain so that background noise is barely audible. In this condition, the high level of reverberation received soon after transmission of a pulse will not overload the receiver for so long a time; thus, echoes can be heard at shorter ranges than would otherwise be possible. Similar remarks apply to nonauditory detection. Finally, when the peak factor of the noise is very large compared with changes of amplitude produced by introducing a primaudible signal, that is, when the total "swing" is large, it may be helpful to use some limiting if this is done with caution.

The fact that a limiter (or other nonlinear system) changes the envelope of the transmitted disturbance means that the composition of the output spectrum of the sound will be different from its input spectrum. Thus, the effects of distortion associated with limiting depend upon system band width, frequency response (gain at various frequencies), operating point and extent of swing. It should be noted that such factors interact with each other; in other words, the effect of varying two of them jointly is not necessarily equal to the sum of the effects produced by independent variation. Consequently, it is probably best to study the effects of such changes by varying two or more factors from test to test and to examine the more important combinations of the various factors.

In a small number of tests conducted at UCDWR, moderate and extreme degrees of limiting were studied for their effects on the primaudibility of echoes masked by noise and reverberation; for the results of the latter tests, see Section 9.2.2. The primaudibility tests with noise backgrounds were made in the usual way, for given signal and noise pairs. Masking tests were made first without limiting and then with either of two degrees of limiting. The test apparatus, minus the distorting networks, showed no more than 2 to 3 per cent of distortion at any gain setting in the useful operating range.

The first type of limiting involved only a gradual flattening of the tops of the positive half cycles of an oscillator tone transmitted by the apparatus, the amount of flattening being somewhat dependent on the signal level. The second type of limiter had a very marked rectifying action, allowing somewhat less than ½ cycle to be transmitted out of each full cycle of a sinusoidal input wave.

Recorded 800-cycle beam-aspect echoes of about 100 milliseconds duration were used as signals, and wide-band thermal noise constituted the masking background. The levels were measured at the input to the distorting network, in the manner already described for recorded echo and noise samples (see Section 8.4.6). They were also measured at the output. Recognition differentials computed on the basis of input and output levels were very nearly equal. The output values are reported in the present discussion.

The signal-background mixture was passed through either of two filters which defined the listening band. The first of these transmitted components between 565 and 1,130 cycles; the second, components between 0.1 and 9 kilocycles. In some tests, the signal-background mixtures were presented to the observers through headphones, and in others by means of loudspeakers, both of which were of high quality. Both distorting networks changed the quality of the presented sounds. Signal and noise became "wheezy," and more so for the severer condition of distortion. Observers expressed annoyance after listening to distorted sounds.

Preliminary results indicate that signal recognition was not significantly impaired for the moderate degree of distortion; and this was true for either loudspeaker or headphone presentation and for both band widths studied. When the extreme degree of distortion was used, with either phones or speaker, recognition suffered by 1.0 decibel in the case of the 565- to 1,130-cycle filter and by 2.6 decibels for the 0.1-9 kilocycle filter. These effects exceeded the variability from all other sources and are regarded as statistically significant. The reasons for this variation with filter width are not clear, although changes in overall gain, in loudness level, and in the number of intermodulation frequencies passed by the filter may all have been important. It seems clear, however, that the types of distortion used in these tests (which probably exceed anything likely to occur in practice) do not seem to be seriously harmful.

8.5 UCDWR TESTS ON REPEATED PULSES

In the development of acoustic fathometers for use on submarines, it was suggested that the use of rapidly repeated pulses might offer some advantage. One of the objectives in fathometer design is to use a signal which an antisubmarine vessel would not be likely to detect, but which would, after reflection from the bottom, yield a recognizable signal at the submarine. To evaluate the possible usefulness of repeated pulses in these and other applications, a program of psychoacoustic tests was undertaken by UCDWR. The first factor of interest was the recognition differential for a string of pulses as a function of pulse length and pulse repetition frequency. Other factors which also affect the recognition differential and which were studied in this program, were the heterodyne frequency, the frequency composition of the repeated pulses, the width of the listening band, and the time-amplitude pattern of the masking noise background.

8.5.1 General Procedure

To generate tonal pulses, a sustained oscillator tone was subdivided into segments of known duration by means of a gating circuit in the signal channel. With this method of pulse generation, a 1-millisecond pulse repeated at the rate of 1,000 pulses per second would produce a sustained tone at the output; in other words, each 1-millisecond pulse would connect into the succeeding one without shift in phase and almost without transient.

The pulses studied were 1, 2, 5, and 9 milliseconds long. The 1-millisecond pulses were studied at a variety of regular pulse repetition frequencies (prf) which could be adjusted from a rate of 4 to a rate of 600 pulses per second. Much of the work was done at a prf of 5.6 per second. For this prf, the interval between pulses is about 0.18 second and is therefore of sufficient duration to permit the ear to resolve the individual pulses. The effect of irregular pulse repetition was also examined. In this case the test administrator could control the occurrence but not the duration of the pulses by tapping a well-adjusted telegraph key. In this way, strings of irregularly spaced pulses could be produced, the spacing between successive pulses varying from an interval of about 1 second to intervals which were as short as the administrator could make them.

Signal level was defined in all cases as the peak level of the individual pulse or pulses used. When repeated pulses were investigated, each signal consisted of a string of pulses extending over an interval of 3 to 5 seconds, during which the pulse level was held fixed.

To study the effect of heterodyne frequency, a number of oscillator tones, varying in frequency between 200 cycles and 6,000 cycles, were fed into the pulser. In all cases in which a pulse length of 1 millisecond and an oscillator frequency less than 1 kilocycle was used, the pulsed signals consisted of less than one complete cycle. This factor, however, appears to have had little if any effect upon pulse detectability. In addition to tonal pulses, pulsed thermal noise was studied. The means employed to produce thermal noise pulses are described later in this section.

All the masking background noises used in these tests were recorded on film. These backgrounds were obtained from four separate sources. Three of the recorded noises had power level spectra given by curve A in Figure 17;

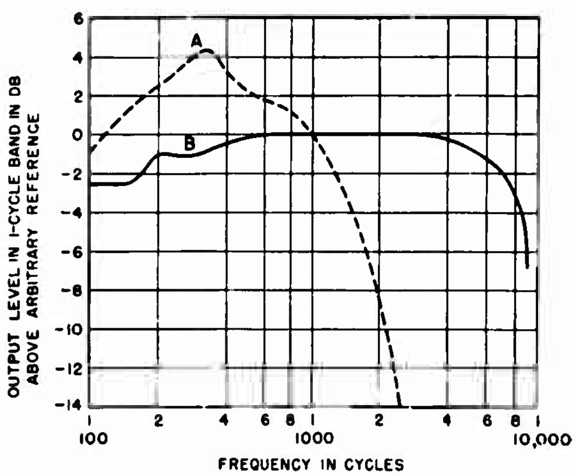


FIGURE 17. Spectra of two types of background used in masking repeated pulses.

the remaining one had a power level spectrum given by curve B in that figure. The three backgrounds of relatively low frequency were recordings of heterodyned self-noise received in the supersonic gear of an S-class submarine; hence the shapes of all three spectra were very nearly identical, being determined essentially by the response characteristics of the receiving gear and of the recording and reproducing systems. One of the three self-noise samples was obtained while the submarine was underway at a depth of 90 feet and a speed of 3 knots. The quality of this noise background was described as "smooth"; it apparently consisted almost entirely of ambient water noise. The second of the self-noise samples was obtained with the submarine moving on the surface at a speed of 10 knots, with the hydrophone axis at a relative bearing of 270 degrees. The quality of this noise was described as "rough," and consisted of scratches superposed upon a hissing sound of variable loudness. For short periods, only the hissing sound could be heard. The third sample of recorded self-noise was obtained while the submarine was lying-to on the surface, charging its batteries with diesels, and with most of its auxiliaries secured. This noise was also described as "smooth," being fairly free of sharp impacts and consisting mostly of ambient water noise. This last sample of recorded self-noise was used in the signal channel only and in those cases in which pulsed noise served as the signal to be recognized.

The fourth background used was a film recording of the output from a thermal noise generator. The spectrum of this sound, as measured at the input to the headphones, is represented by curve B in Figure 17; its character was smooth. The shape of its spectrum is largely a reflection of the band widths of the noise channel and the mixing stage. Since the spectra of pulses were restricted in much the same way by the band width of the test apparatus, it was unnecessary to use filters after the mixing stage (for the sake of confining the pulse spectrum to the limits of the masking noise band) when this wide-band thermal noise sample was used as the masking background. However, this statement does not apply to studies in which the masking noise sample had the spectrum described by curve A; some of the anomalous results described below, in connection with tests in which the masking background was confined to the narrow band, seem to have been produced by the fact that the pulse spectrum extended beyond that of the masking noise. The wide-band thermal noise represented by curve B was also used in some of the tests in which noise pulses were studied. For such tests, two separate sections of the noise film recording were used in the background and signal channels, so that there was no correlation between the time-amplitude patterns of the pulses and the masking background. The specific type of noise used in any case is indicated in the figures showing the results of the listening tests.

Several properties of the presented signals and backgrounds may have influenced these tests. The deviations from absolute flatness shown by curve *B* may have affected, to some extent, the results obtained with pulses of different intrinsic frequencies. Also, while the headphones used were of relatively high quality, they probably had a somewhat poorer response in the high-frequency region than in the low. Finally, it seems probable that, when the wide-band noise depicted by curve *B* was presented to the listeners at a comfortable loudness level, the components below about 200 cycles did not reach threshold (see Figure 79 in Chapter 4); hence, the results of masking tests with pulses of very low intrinsic frequency may have been affected in part by the fact that the signals were threshold rather than masking limited.

Some of the tests were made with several observers participating. In the remainder, the results were obtained by a single observer in self-administered tests. This observer had participated in the group test as well, and his responses were found to be reliable and very nearly typical of the group response. Two types of test were conducted. In one, the observers had no foreknowledge of the character and quality of the sound which was to be identified as the signal. As indicated below, results in this type of test were significantly different from the results obtained when the observers knew the nature of the signal in advance.

In the various figures given in this section, experimental points are shown only for the one case in which they are available (Figure 18). The report describing the results of these tests[6] states that in no case were the experimental points more than 2 decibles away from the line drawn.

8.5.2 Effect of Pulse Repetition Frequency

The effect of pulse repetition frequency on the audibility of 1-millisecond pulses is shown in Figure 18. In this case, pulses were repeated at regular rates shown on the horizontal scale in the figure. The ordinates represent recognition differentials relative to the standard reference band (0.1 to 10 kilocycles) corrected to a flat background spectrum. A glance at the observed spectrum of noise *B* indicates that, in fact, the spectrum of the masking background was not flat and that the equivalent rectangular noise band is much more nearly included between the limits 0.2 and 6 kilocycles. Actual integration of the total power contained in noise

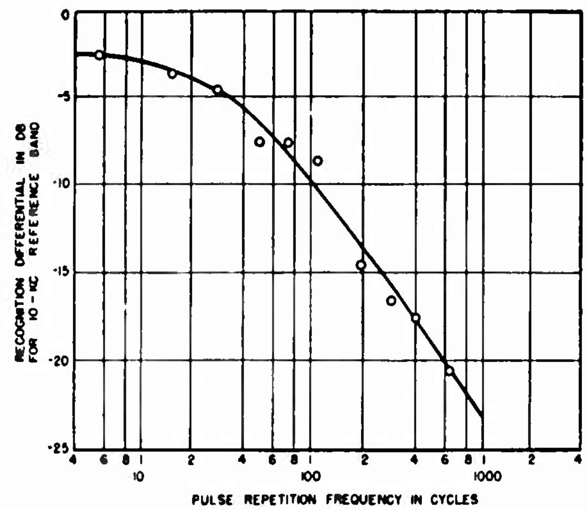


FIGURE 18. Effect of pulse repetition frequency on recognition differentials. The signal consisted of 1-millisecond 800-cycle pulses repeated at the rates indicated on the horizontal scale. The background was wide-band thermal noise. For comparison, the recognition differential for a single presentation of a 1-millisecond 800-cycle pulse masked by a 10-kilocycle noise band is 6.5 decibels (see Figure 8).

B shows that the equivalent flat spectrum has a width of about 5,650 cycles. Hence the observed values of the signal-to-noise ratio at primaudibility have been diminished by 2.5 decibels to convert them to recognition differentials for a flat noise background in the standard reference band. This correction is somewhat uncertain since the meaning of the data reported in reference 6 is not wholly clear.

The circles in Figure 18 represent the observed points; the heavy line has been drawn to fit these points as well as possible in a smooth manner. Figure 18 indicates that the audibility of multiple pulses does not change by more than 1 decibel as the pulse repetition frequency increases from 5 to 20 per second. With increasing rates of repetition the audibility improves; for a just audible string of pulses, the required signal intensity decreases by 10 to 13 decibels for a tenfold increase in pulse repetition frequency. This result may be quantitatively explained as a result simply of the increase in average signal level with increasing prf, if the response of the ear is assumed to depend primarily on the total energy in the signal. Since the average signal energy is proportional to the number of pulses per second, this energy presented to the ear will increase 10 decibels per tenfold increase of prf. The conclusion that the masking of a sufficiently short pulse or group of pulses is determined primarily by the total signal energy presented during the ear's integration time is in general agreement with the results already discussed and summarized in Section 8.1.

One would expect that as the time interval between pulses approached zero and the pulses merged to form a sustained tone, the observed recognition differential would approach that observed for a sustained tone masked by wide-band noise. Since the pulse length used for the data in Figure 18 was 1 millisecond, the *RD* for a prf of 1,000 per second should equal the *RD* found with the sustained tone. Comparison between these two recognition differentials shows, in fact, very good agreement between them. The repeated pulse *RD* for a prf of 1,000 per second, found by extrapolation in Figure 18, is -23 decibels, referred to a flat noise band 10 kilocycles wide. Since the critical band for an 800-cycle tone is 38 cycles wide, the computed recognition differential is 10 log 38/10,000 or -24.2 decibels. The slight discrepancy between these two results may be due to the fact that, when an 800-cycle tone is subdivided into 1-millisecond pulses, the period of the generated pulses is much more nearly that of a 1-kilocycle wave than it is of an 800-cycle wave. If this type of distortion actually played a part in present tests, a critical band width of 50 cycles should be assumed in the preceding calculation, in which case the computed *RD* is in exact agreement with that derived from Figure 18.

There is, however, another way of looking at this problem. Mathematically, a string of repeated pulses can be analyzed into a spectrum of sustained tones whose sum is equivalent to the string of pulses and is physically indistinguishable from that string. It might be expected that such a string of pulses would be audible when one or more of the tones in the spectrum was equal in intensity to the noise level in the corresponding critical band.

This type of analysis may be expected to be relevant to the ear only for values of prf greater than about 20 per second. With more widely spaced pulses, the ear will resolve the separate pulses and respond to them as units. For prf's more than 20 per second, the sound has a high degree of roughness, but the successive pulses are not heard as clearly defined individual units. For the higher values it is therefore relevant to investigate the intensities of the individual tones making up the spectrum of the repeated pulses and to compare these with the masking background.

Unfortunately, the exact analysis of the strings used in these tests is somewhat complicated, since the square-wave modulation introduced by the gating circuit bears no simple relationship to the phase of the oscillator tone. With certain approximations, an analysis can be carried out, but the results are not easily reconciled with the observed masking data. In fact, the slope found on this approximate theory is roughly a 20-decibel decrease of recognition differential for each tenfold increase of prf, since the intensities of the tones in the equivalent spectrum are proportional to the square of the prf. It is possible, however, that, because of their harmonic and phase relationships, the components of such a spectrum are integrated so that they become primaudible as a group. Since the number in such a group would be proportional to the prf, the observed slope of 10 decibels per decade would be expected. More detailed study is required to indicate the relationships between the performance of the ear and the spectrum of this type of signal.

8.5.3 Effect of Pulse Type

Observations were made of the masking of repeated pulses of various types with a pulse repetition frequency of 5.6 per second. Among the variables that were studied in this connection were the pulse length and the frequency of the tonal pulse used. In addition, the recognition of noise pulses was studied.

Figure 19 shows the recognition differentials obtained with pulses of lengths between 1 and

10 milliseconds and with frequencies of 200 to 6,000 cycles. No significant difference in recognition differentials, that is differences exceeding the experimental error of 1 to 2 decibels, appears to have been obtained by varying the heterodyne frequency between 200 and

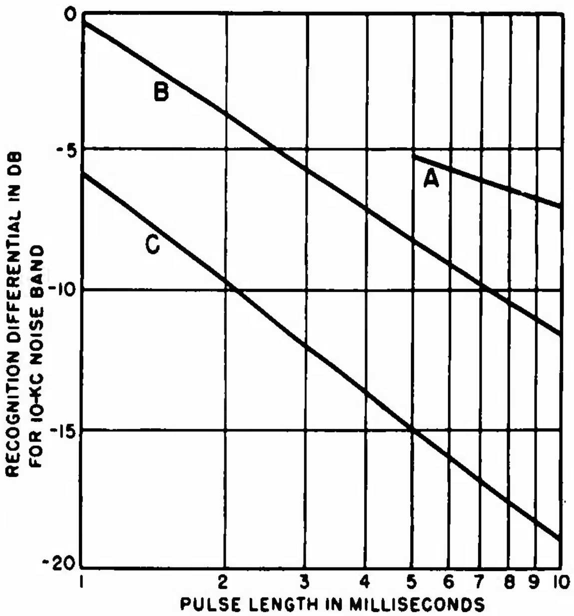


FIGURE 19. Effect of pulse length on audibility of pulses repeated 5.6 times per second. Curve A is for 200-cycle pulses masked by noise whose spectrum is given by curve B in Figure 17. Curve B is for pulses with frequencies between 0.5 and 6 kilocycles masked by noise (curve B in Figure 17). Curve C is for 700-cycle pulses masked by noise (curve A in Figure 17).

6,000 cycles, although the 200-cycle data in the figure indicate a loss of detectability amounting to 3 to 5 decibels, and the 700-cycle data imply a gain of 6 to 7 decibels.

It will be noted that the slope of curve B shown for the pulses of 0.5 to 6 kilocycles, masked by wide-band thermal noise, is 11 decibels for each tenfold increase in pulse length. This is to be compared to the slope of 10 decibels to be expected if the response of the ear to the short pulses is determined primarily by the total energy presented to the ear per second. This relatively close agreement provides added confirmation that the ear does, in fact, behave as an integrating mechanism for the masking of short pulses; it has already been noted in Sections 8.1.2 and 8.4.4, however, that this is not in agreement with the behavior of the ear in estimating the loudness of a short pulse, and this apparent difference between the results of the loudness and the masking studies still remains to be explained.

The data shown in curve A for a pulse frequency of 200 cycles and a masking background of wide-band thermal noise are based on two pulse lengths, 5 and 9 milliseconds respectively. Shorter pulse lengths were not used since it was felt that these would contain too small a fraction of a cycle. The discrepancy between curves A and B is probably not real. For example, when a wide band of flat thermal noise is presented at a comfortable listening level, the noise components at frequencies of 200 cycles or lower are likely to fall below the absolute audibility threshold (see Figure 79 in Chapter 4). More data would be required, however, to give reliable results on this point.

For comparison with the data on single pulses, the curves in Figure 19 have been transcribed to Figures 8 and 12. To express the observed recognition differentials in terms of a reference band 1 kilocycle in width, 10 decibels have been added to the ordinates shown in Figure 19. The recognition differentials found in this way are then comparable to those discussed in Section 8.4. It will be noted that this curve lies about 6 decibels below that for the single pulses, but that its slope is comparable. Thus, it may be inferred that a repetition rate of 5.6 per second gives an RD 6 decibels lower (more favorable) than a single pulse of the same length. This may be compared with a gain of 3 decibels found in Section 8.1.3 from a single repetition of a pulse.

For comparison with curve B, curve C in Figure 19 applies to 700-cycle pulses masked by rough self-noise with the spectrum given by curve A in Figure 17. The spectrum of the pulsed signal extended beyond the limits of the rather narrow noise band; this is to be expected theoretically and was verified by narrow-band filter analyses of the signal spectra. Thus, detectability in the case of curve C was better than that in the case of curve B, because the pulse spectrum was wider than that of the noise. The results are not representative of any practical situation, in which the same receiving system would be used for both signal and noise. This observation illustrates the danger inherent in tests in which the signal and noise are not passed through a common filter. It also indicates that, even for short pulses, the ear continues to act as a frequency analyzer and the masking effect of wide-band noise comes primarily from those noise components with frequencies in the immediate neighborhood of the signal components.

The results obtained in the study of thermal noise pulses are summarized in Figure 20. In

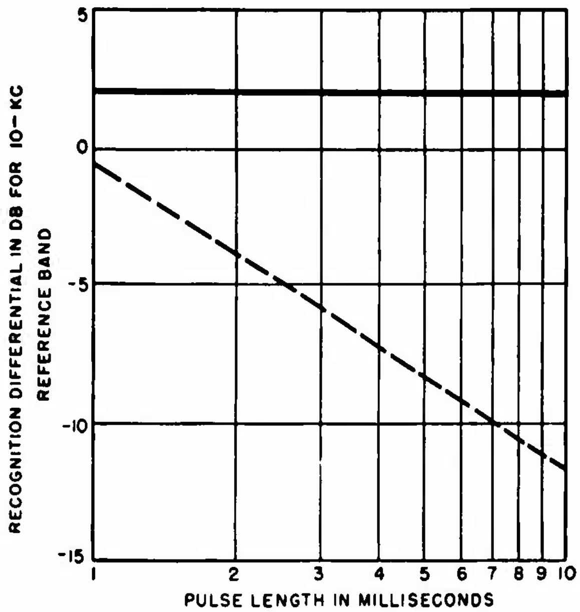


FIGURE 20. Audibility of noise pulses (curve *B* in Figure 17), masked by a noise background (also curve *B* in Figure 17). For comparison, the dashed curve, taken from Figure 19, shows recognition differentials for tonal pulses with frequencies between 0.5 and 6 kilocycles. The pulse repetition frequency in both cases with 5.6 times per second.

this figure the ordinates show recognition differentials, referred to the standard reference band, for pulses of the indicated duration, repeated 5.6 times per second. The upper line gives the measured values for pulses of thermal noise, while the lower, transcribed from Figure 19, represents the data for tonal pulses of frequencies between 500 and 6,000 cycles. It will be noted that the upper line shows a recognition differential independent of pulse duration in the range from 1 to 10 milliseconds. The two lines, if extended to the left, would intersect at a pulse length of about 0.6 millisecond.

Although the constant recognition differentials found for the thermal noise pulses are perhaps surprising, it is to be expected that this curve should show a slope less steep than that for the tonal pulses. It is evident, however, that the two curves should approach equality for sufficiently short pulses. This result follows either from the increasingly wide spectrum of a shorter and shorter pulse or from the observed fact that a very short pulse has no apparent tonality. Since a very short tonal pulse is indistinguishable from a very short noise pulse, the recognition differentials shown by the curves in Figure 20 should approach equality for short pulse lengths. On the other hand, the solid line should lie above the dashed when the pulse length is much increased. This follows from the fact that a distributed noise is less audible than a pure tone in the presence of a wide-band noise (compare the wide-band recognition differentials in Figures 58 and 59 of Chapter 4). While a sustained tone at 1,000 cycles can be heard 23 decibels below a flat background 10 kilocycles wide, a noise similar in spectrum to the background and with no modulation could be at most 3 decibels below the background at primaudibility. Thus, the two lines in Figure 20 would be expected to diverge with increasing pulse length, the difference between them approaching a maximum value of 20 decibels. The behavior shown in Figure 20 therefore appears reasonable.

Data on masking of noise pulses by noise are also given in the next section for the situation where the narrow-band noise source shown in curve *A* of Figures 17 was used for both signal and background. These data give quite different results from those obtained with the wider band noise, and in particular show recognition differentials decreasing with increasing pulse length (see Figures 21 and 22). However, these tests with narrow-band noise pulses are not too reliable, since pulsing the noise may have widened its spectrum, giving audible components at frequencies where the unpulsed noise background was weak. Hence, only the data on wide-band thermal noise will be considered at this point.

Although these data are neither sufficiently accurate nor measured over a sufficiently great range of pulse length to allow precise conclusions, comparisons of these findings with other results are of interest. Since the spectra of signal and background in these thermal noise studies are nearly identical, the recognition differentials shown by the solid line in Figure 20 give the ratio of signal to noise in each critical band at primaudibility. They may therefore be compared with the continuous curve in Figure 11 which gives the signal-to-noise ratio in a critical band for a tonal pulse at primaudibility. This curve has been computed theoretically from the observed dashed curve, taking into account the increasing width of the pulse spectrum with decreasing pulse length. It is perhaps significant that the computed critical-band *RD* increases by less than 2 decibels as the pulse length decreases from 10 milliseconds to 1 millisecond. This agrees, to within the experimental error, with the constant recognition differentials shown by the upper line of Figure 20. Thus, it is possible that in this range of pulse lengths the critical band *RD* actually changes very little. Additional evidence for this tentative point of view is presented in Section 9.2.2.

Studies were also made of the effects produced by passing the signal through filters of different widths. In all cases, however, the filter widths were at least as great as the reciprocal of the pulse length, and therefore passed most of the signal energy. While some distortion of signal envelope was produced, this effect should not be aurally noticeable at the short pulse lengths used. In agreement with expectation, the standard-band recognition differentials found with filters were found to be equal to those obtained without filters, to within the experimental error. Since these recognition differentials are expressed in terms of a wide masking band it follows that the primaudible signal level was unaffected by the filters used, although the level of the actual noise presented was of course affected by the filters.

8.5.4 Effect of Irregular Repetition Rate and Amplitude Modulation

Figures 21, 22, and 23 show the effects of irregular as compared with regular rates of pulse repetition. The effect of this factor is shown for pulses of the indicated durations and frequency compositions and in the presence of a noise background with the spectrum shown by curve *A* in Figure 17. Both rough and smooth noise backgrounds were used. As indicated in the preceding section, this background noise did not adequately mask the pulses. Consequently, the ordinates in Figures 21 through 23 are

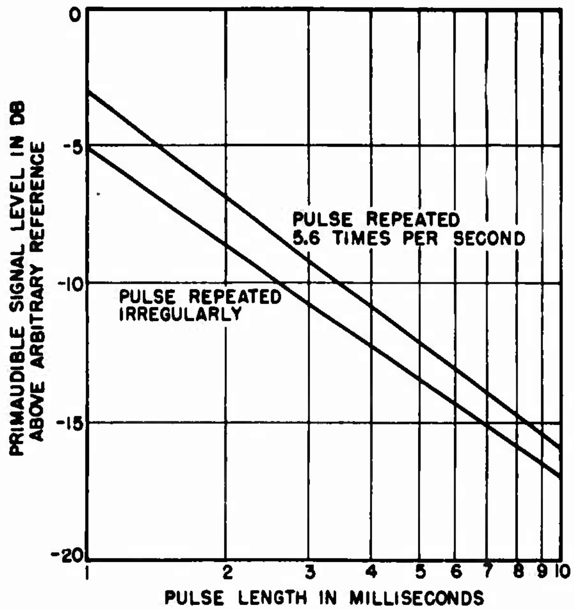


FIGURE 21. Effect of random-interval pulsing for a 700-cycle tone, masked by rough self-noise (curve *A* in Figure 17).

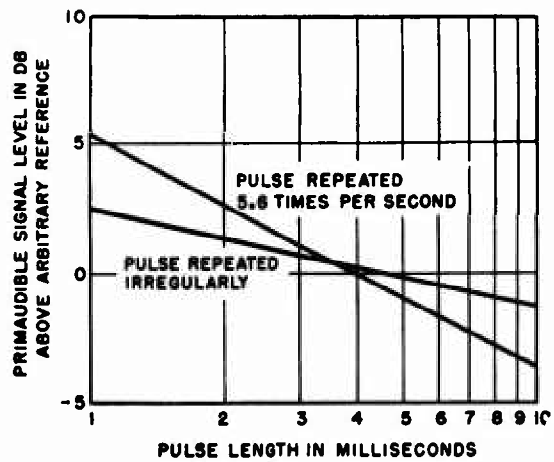


FIGURE 22. Effect of random-interval pulsing for smooth self-noise, masked by rough self-noise. Spectra of signal and background are given by curve *A* in Figure 17.

arbitrary and do not represent recognition differentials. The results shown in the various graphs may, however, be used for comparing the effects of regular and irregular pulse repetition rates, and also smooth with rough background noise types. Due to the spreading of the pulse spectra beyond the limits of the masking noise band, even these comparisons may not be truly representative of the results which would be obtained in the field.

Figure 21 applies to a 700-cycle tone of the indicated duration pulsed at a regular rate of 5.6 per second, and also pulsed in an irregular fashion. The apparent advantage of some 2 decibels for irregular as compared with regular pulsing may be due to the fact that the rough self-noise was observed to have an approximately cyclical time-amplitude pattern of its own. Possibly, therefore, the irregular pulsing was easier to detect than the regular because it could be more readily distinguished from the time pattern exhibited by the noise background. In other words, when hand keying is used, a signal rhythm should be selected which is as different from that of the noise as possible. It is also possible that differences in the average pulse repetition frequency may be in part responsible for the difference between these two curves.

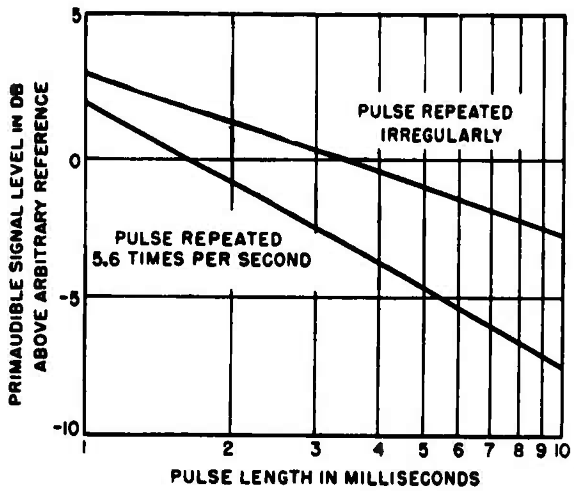


FIGURE 23. Effect of random-interval pulsing for smooth self-noise pulses, masked by a smooth self-noise background. Spectra of signal and background are given by curve *A* in Figure 17.

Figure 22 depicts changes in performance due to regularity of pulsing for the case in which a signal obtained from smooth self-noise was detected in the presence of rough self-noise. Before pulsing, both of these noises had the characteristics shown by curve *A* in Figure 17. The results given in this figure correspond more nearly to expectation than do those in Figure 21, since it might be expected that regular patterns could be more readily recognized with certainty than irregular patterns of the same pulses. In both cases, however, the *RD* for noise pulses decreases with increasing pulse length, in contrast to the data for wide-band noises shown in Figure 20. Similar results were obtained when smooth self-noise was used as the masking background (see Figure 23). This discrepancy between results found with wide-band and narrow-band noise has already been discussed in the preceding section.

The importance of time pattern is illustrated in Figures 24 and 25, which show the relative audibility of signal pulses obtained from smooth self-noise, and masked by either smooth or rough self-noise. In these figures, the curves for a rough noise background have been transcribed from Figure 22, and those for a smooth background from Figure 23. The greater masking efficiency of the rough background noise shown in both cases was apparently associated with its time-amplitude pattern. Thus, when the rough masking noise was used, the observers

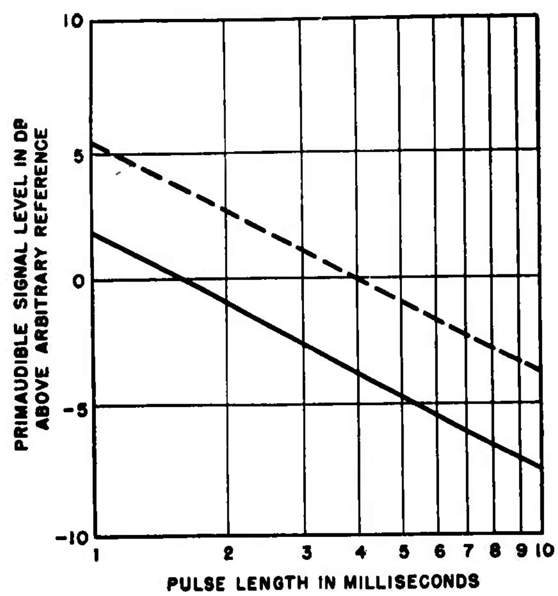


FIGURE 24. Effects produced by the time pattern of masking noise for smooth self-noise repeated 5.6 times per second. The solid curve applies to a smooth self-noise background, while the dashed curve applies to a rough self noise background. Spectra of signal and background are given by curve *A* in Figure 17.

had to listen for several seconds in order to be sure that a signal was present and to detect its rhythm (see also Section 5.5).

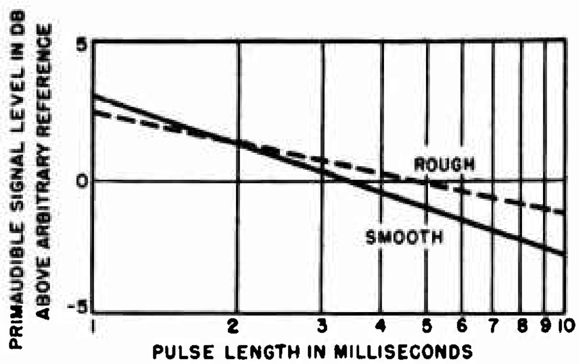


FIGURE 25. Effects produced by the time pattern of masking noise for smooth self-noise pulses repeated at random. The solid curve applies to a smooth self-noise background, while the dashed curve applies to a rough self-noise background. Spectra of signal and background are given by curve A in Figure 17.

Another significant observation should be noted at this point concerning the effect on recognition differentials of advance familiarity with the presented signal. It was generally observed that a just audible signal of unknown character was 3 to 8 decibels higher than a signal of known character when irregular pulsing was used; in the case of regular pulsing, the *RD* for unknown signals was 3 to 5 decibels above that for known signals. The report describing the results[6] indicates that the magnitude of this effect should be considered the same for regular and irregular pulsing, in view of the experimental uncertainties. For comparison, it will be recalled that similar tests with ship signals (see Section 4.1.5) indicate that the primaudible signal-noise ratio for unknown signals is 2 to 3 decibels higher than for familiar signals. It was the opinion of the observers in these tests that strings of very short pulses would not attract attention at an enemy listening location unless they were extremely loud with respect to the background or unless they suddenly increased in loudness as the bearing of the enemy receiver was changed.

Chapter 9

REVERBERATION MASKING OF SIGNALS WITHOUT DOPPLER

REALISTIC TESTS on masking by reverberation are considerably more difficult to perform than on masking by noise, since the characteristics of reverberation are so varied and so complicated. As a result of this complexity of reverberation, it is not feasible to use simulated reverberation in masking tests, since no method of simulation has been devised which produces sound nearly enough identical with observed reverberation. Instead, all studies of this subject have used reverberation background received in sonar gear, either presented directly in a test or recorded for subsequent presentation. The signals used have included injected pulses as well as recorded echoes. In almost all cases, one of the primary difficulties has been the narrow frequency band to which CW reverberation is confined. As a result of this small frequency spread in the masking background, it is not easy to avoid spurious effects resulting from the finite width of the signal spectrum. There are a number of instrumental effects which can broaden the spectrum of a pulse or even of a recorded echo beyond the spectrum of the reverberation, and much of the effort in these tests has been devoted to the elimination of this effect.

9.1 BELL TELEPHONE LABORATORIES TESTS

In a pioneering study,[1] masking experiments were made by BTL with pulses and recordings of real echoes both of which were mixed with recorded CW reverberation samples. The pulse signals were rounded as well as rectangular. Both auditory and visual detection were studied, but only the auditory results are described here.

9.1.1 General Procedure

The reverberation background used in these tests was a playback of disk recordings. Each disk contained about 100 samples of volume reverberation, produced in deep water by a CW transmission 5, 25, or 100 milliseconds long and received at bearings of 90 or 270 degrees by a vessel moving at about 8 knots. The frequency put into the water was 24 kilocycles, and the received reverberation was heterodyned to 800 cycles and transmitted by an FM link to a shore station where it was recorded. Successive reverberations in each recording were spaced 4 to 6 seconds apart.

Owing to overloading of the receiving-recording system, the initial portion of each reverberation was of nearly uniform amplitude and had appreciable harmonic content. This part was followed by a second portion in which the sound level fell more or less rapidly and was featured by rapid fluctuations of amplitude, or blobs, with durations approximately equal to that of the emitted ping. The third section of each recorded reverberation was of relatively low level, approaching that of needle scratch and other noises incidental to the recording and playback. Signals were mixed with these backgrounds in such a way that the signal occurred during the second, strong, but decaying, portion of the reverberation samples. To minimize annoyance and confusion which might be produced by the playback noise associated with the third segment of the reverberation samples, a 3,500-cycle low-pass filter was included in the background channel.

The reverberations were recorded at bearings of 90 or 270 degrees in order to eliminate the effects of own-doppler as much as possible. However, careful examination of recorded reverberation samples indicates fairly large variations in pitch of the reverberation are not uncommon. Such pitch changes may be expected to influence recognition differentials; hence it is desirable to eliminate from masking tests all reverberation samples which show this tendency. While no selection of the reverberation samples for the purpose of eliminating variable pitch items was made in the tests under discussion, comparison with other observations (see Figure 12) implies that instability in pitch of the reverberation background probably played no very large role in determining the overall results obtained in these tests. Further study will be required before it is possible to decide whether systematic changes of reverberation frequency with time are important in practice.

Either pulse or echo signals were injected at the same range in each case; that is, the length of time by which the signal followed the start of each reverberation was fixed. This procedure involves the risk of introducing unrealistic practice effects by conditioning the observers to concentrate on the signal at a particular instant during the course of each reverberation and to disregard misleading blobs which may occur earlier or later in the sample. However, evidence discussed in Section 9.2.2 indicates that the probable magnitude of such an effect in systematically diminishing recognition differentials probably did not exceed 1 to 2 decibels.

The sounds used were presented individually to the members of the test group, usually 4 to 6 in number. Each test contained 100 reverberation samples (100 listening intervals). No more than two such test sequences were presented to any observer at a single sitting; and the factors tested in two such sequences were selected, from among those included in the overall program, in as random and unrelated a manner as possible.

Before the tests were begun, a number of subjects were given instruction and practice in masking tests of the kind under study. All the observers actually used in the study under discussion were drawn from the more able group of the subjects examined, on the assumption that more uniform data would be obtained and that there would be less practice effect to modify the performance of the abler group with time. Later tests indicated that the magnitude of the practice effect was probably less than the variation in performance of a single individual in successive tests. Checks with groups of 5 to 6 observers indicated that the results obtained in a single test given to the group were reproducible to within about 2 decibels.

During each sequence of 100 samples, the level as well as the range of the injected signal was held fixed. The observers could adjust the level of the presented sounds for maximum comfort; thus, the gain settings chosen varied somewhat among different observers. Approximately 50 signals were injected at random in each sequence of 100 listening intervals. In successive sequences, the gain in the background channel was left unchanged at a constant value, and the level of the injected pulse was varied in steps of 2 to 3 decibels from a value producing about 95 per cent recognition probability to a value giving approximately 50 per cent recognition probability. The highest levels were used first, in order to minimize practice effects and to establish listener confidence. In addition, competition was fostered among the observers in order to relieve the monotony of the task. Just before each test, the observers were permitted to listen to as many signal-reverberation mixtures as they desired and at the same signal level to be used in the forthcoming test. In the report describing these tests,[1] it is stated that "for some of the more difficult tests [that is, lower signal levels], and at the request of the observers, the practice was begun with a higher signal level which was gradually reduced to the test value. This procedure often improved considerably the observers' ability to detect the more obscure signals. If, after practice, the observer felt that for any reason his performance was not up to par, testing with that individual was postponed. The observer was also made to feel free to stop testing at any time he desired. While these privileges were seldom exercised, they helped to relieve any tendency toward nervous strain and thus contributed to better and more nearly uniform results."

General observations made in reverberation masking programs indicate that the cues are often elusive and that experience and favorable test conditions do have an important influence on the character of the results obtained. It should be noted, however, that such test procedures tend to define the upper limit of performance and may not be directly applicable to the more complicated and less favorable set of conditions typically met in practice.

All tests were conducted in a soundproof room, and loudspeaker presentation was used exclusively. One of the loudspeakers used was a regulation Navy model; the other was of higher quality, having good response over a broader band than the first and passing more of the low-frequency sounds, that is, it was less sharply tuned than the first.

The listeners expressed their judgments of signal audibility by pressing a telegraph key, which actuated recording circuits designed to count separately all errors of omission and commission. Failure to depress the key within a 2-second interval beginning with the signal was scored as an omission, while a vote at any other time was registered as a commissive error.

The signal-to-reverberation ratio in any test was defined as the ratio of the average instantaneous power of a signal to that of the reverberation during a period equal in length to that of the signal and coincident with it. When the signal was a pulse, its power was readily measured; the method of measuring the power of recorded echoes is described in Section 8.4.6. Since the reverberation level (at the instant the signal was injected) varied from one background sample to the next, a statistical average of the 100 reverberations in each test was used for the purpose of defining reverberation power. To do this, the power-time trace for every fifth reverberation on a disk (20 in all, for each disk) was obtained with a power level recorder whose rated writing speed was 360 decibels per second. The power measurements for the 20 samples selected were averaged at ¼-second intervals, and the arithmetical means of these individual readings were used in computing signal-to-reverberation ratios.[a] The standard deviations from such averages varied somewhat from disk to disk, ranging from 1 to 3 decibels. This implies that the extremes of fluctuation extended over a total range of about 18 decibels.

The finite response times of power level recorders tends to smooth out the extremes of fluctuation because the instrument shows the integrated power fed to its terminals during its integration time. The effect upon calculated recognition differentials of various methods of measuring the level of reverberation background is discussed in some detail in Section 9.2.2, where it is indicated that the method just described is perhaps as satisfactory as any when the purpose of the measurements is to obtain predictions of probably performance in the field. The report describing the tests discussed in the present section[1] assigns an estimated value of less than 1 decibel to the net effect of all sources of objective error, such as faulty calibration or instability in the test apparatus.

9.1.2 Recognition Differentials for Injected Pulses

In the first set of tests, pulses were generated by an 800-cycle oscillator in the signal channel. The harmonic content of this oscillator tone was more than 40 decibels below the fundamental. A gate circuit, or electronic relay, was used to generate rectangular pulses 5, 25, and 100 milliseconds long. These pulses displayed no noticeable switching transient, and, in the masking tests, were mixed with reverberations produced by transmissions of the same duration as that of the pulse signal. The evidence discussed in this section indicates that the envelopes of these artificial pulses probably had somewhat squarer corners than those used to produce the reverberations, that is, the signal spectra contained stronger side-band frequencies than did the background spectra.

Rectangular Pulses

The results obtained in masking tests with extremely square-cornered pulses are listed in Table 1. The tabulated percentages refer to the fractional number of injected signals which were correctly identified. Commissive errors were nearly negligible, averaging about 3 per cent in all tests to which reference is made in this table, and such errors showed no significant dependence on pulse length or signal-to-reverberation ratio. It is generally agreed that the results quoted in Table 1 are anomalous because the test conditions deviated from those which exist in the field. These data are nevertheless worth examining because they help to

[a] For purposes of comparison, the average power level (the geometrical mean) was also computed and was found not to differ significantly from the average power.

define the factors involved in the auditory detection of echoes masked by reverberation.

The results obtained in these tests are atypical in a number of ways. Thus, as shown by Table 1, performance was as good for the 5-millisecond pulse as for the 100-millisecond; in fact, ability to detect the signal is seen to be improved slightly by diminishing pulse duration. In addition, Table 1 implies that tonal pulses can easily be recognized when they are more than 12 decibels below the level of the reverberation centered at the same frequency. Such a finding would not be surprising for the masking of one sustained tone by another of nearly equal frequency, because of the assistance rendered to the ear by the occurrence of beats. It is obvious, however, that if as many as 3 beats occur within the active duration of a 5-millisecond pulse, the rate of beating would be about 600 per second. Such rapid beats are not very audible (see Figure 16 in Chapter 2). Furthermore, this would imply that the intrinsic frequencies of the pulse and the reverberation differed by about 600 cycles. Observations made during these tests, however, show that the intrinsic frequencies of the pulses and the recorded reverberation samples were nearly identical.

TABLE 1. Aural recognition of rectangular 800-cycle pulses.

Pulse-to-reverberation ratio in db	Recognition probability in per cent					
	100-millisecond pulse		25-millisecond pulse		5-millisecond pulse	
	High-quality speaker	Service model	High-quality speaker	Service model	High-quality speaker	Service model
−2	97	96	90.5	89	97	93
−7	98	84	83	82	96.5	87
−12	71.5	60	73.5	75	79	67.5

It is therefore believed that the source of the high degree of signal detectability in these tests was the fact that the signal pulse was much squarer than that used to produce the reverberation. Thus, the pulse spectrum contained stronger side bands than did the reverberation spectrum. Consequently, the pulse could be heard at such surprisingly low levels because its spectrum extended beyond the masked region. In practice, the frequency compositions of reverberation and of an echo without doppler are probably closely similar; hence, it does not seem possible to exploit this observation in order to improve performance when reverberation is limiting in the field.

The interpretation of the rectangular pulse data just outlined is consistent with the observation that the subjects could hear a sharp hiss or click whenever the square pulse was injected, even at very low signal-to-reverberation ratios. The effect of the disparity between the widths of pulse and reverberation spectra was intentionally emphasized in order to test this explanation. This was done by placing a 140-cycle band-pass filter in the reverberation channel, in order to restrict further the width of its spectrum. Under these circumstances, the observers were able to detect a 100-millisecond pulse when its level was 26 decibels below the level of the 100-millisecond reverberation background. These results are consistent with the previously expressed view that the ear continues to act as an analyzer even for sounds of very short duration and may also explain why the performance with the high-quality loudspeaker was somewhat better than with the standard Navy speaker. Since all of the other tests in the series under discussion were conducted with a high-quality loudspeaker exclusively, they do not indicate whether the response of speaker or headphone makes a significant difference in the practical case (when signal and reverberation have spectra of nearly equal widths). The very gradual slope of the transition curves partly defined by Table 1 is probably related to the multiplicity of cues (loudness, frequency, envelope, and others), so that the observers did not use the same standard of judgment at all signal levels.

Rounded Pulses

In order to eliminate the spurious cues supplied by the strong side bands in the spectra of the rectangular pulses, the signal-reverberation mixture was transmitted through a band-pass filter inserted between the mixer and the high-quality loudspeaker. The results obtained in this way are in fairly good agreement with those of other tests (see Figure 12) and probably furnish a reliable guide to the performance which may be expected in practice.

The band-pass filter used was centered at 765 cycles. Frequencies 35 cycles above or below this midfrequency were attenuated to the extent of 2 decibels by the filter, and frequencies removed 70 cycles from the midpoint were attenuated by 10 decibels. Since pulse and reverberation spectra were centered at 800 cycles, they did not coincide exactly with the midpoint of the filter pass band. Pulse and reverberation levels were measured in the signal and background channels. It would probably have been preferable to make these measurements at the filter output, but the available evidence suggests that the recognition differentials obtained would not have been substantially modified by this change. The effect of the filter in rounding the envelopes of both signal and reverberation is illustrated in Figure 1.

Pulses both 100 and 25 milliseconds long were studied with the aid of the band-pass filter. The 5-millisecond tests could not be repeated, since the essential spectrum of such a short pulse is 400 cycles wide and thus exceeds the pass band of the filter. It is possible, also, that the 25-millisecond data obtained with the filter are somewhat less reliable than are the corresponding 100-millisecond data. In the former case, the essential spectra of signal and background are 80 cycles wide; since the midfrequency of these sounds was approximately 35 cycles above the midpoint of the filter, the envelopes of the 25-millisecond pulses may have been excessively distorted.

The results obtained in the masking tests with the filter are shown in Figure 2. The upper thin lines represent the relative number of pulses correctly identified. Since use of the filter increased the resemblance between signal and background and since the observers were

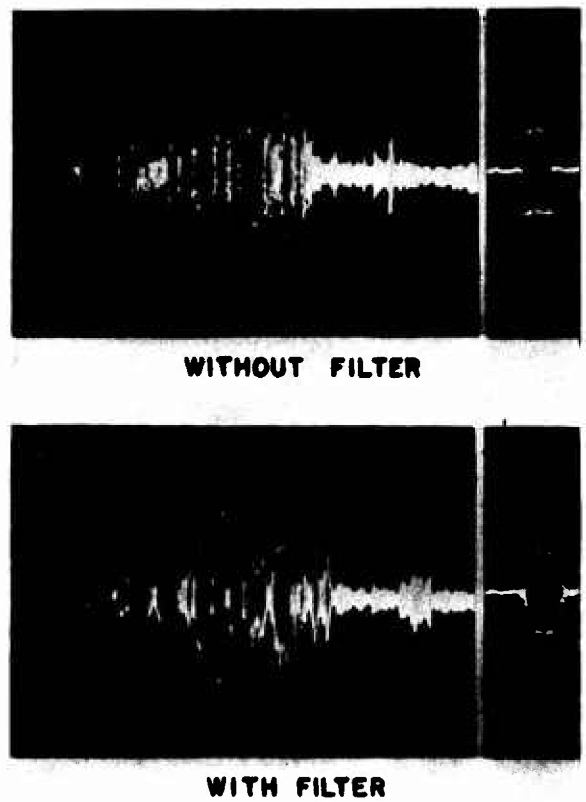


FIGURE 1. Effect of the filter on pulse and reverberation envelopes. Time scales of the pulse envelopes, at the right in both cases, are expanded relative to those of backgrounds, which are shown at the left.

encouraged to adopt a responsive rather than a cautious attitude, the number of commissive errors increased fairly rapidly with diminishing signal-to-reverberation ratio, approaching 15 to 20 per cent for the lowest ratios shown in Figure 2. In the practical case observers are usually trained to be cautious, with the result that commissive errors are generally much less. The observed recognition probabilities have been corrected for this effect. The correction used here is, of course, only an approximation to the actual situation but probably represents the essential features of the problem.

To derive this correction, the recognition probabilities for a cautious and uncautious observer may be considered, when each is presented with signals of constant level in the presence of successive reverberations. Let p_u be the recognition probability for the uncautious observer, defined as the number of times that a presented signal is reported heard, divided by the number of times the signal is presented, and multiplied by 100 to give a percentage. Let p_c be the corresponding quantity for the cautious observer. When a blank interval is presented, the cautious observer will in general report no signal. The uncautious observer will, however, report a signal for q per cent of the blank presentations.

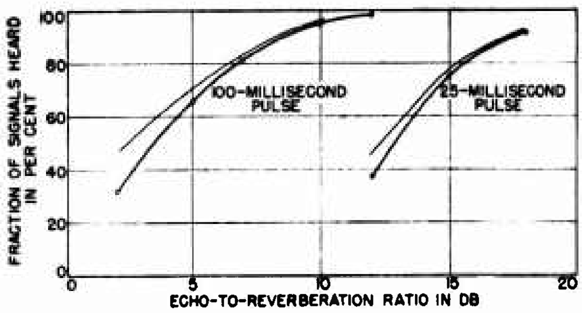


FIGURE 2. Detection probability for 25- and 100-millisecond pulses masked by reverberation. Upper curves give raw scores; lower curves show the points, corrected for guessing.

To obtain a relation between p_c, p_u, and q, the observed fact may be used that a signal which is reported by a cautious observer will usually also be reported by the uncautious observer. Thus, the difference between the two observers will be that the uncautious one will claim to have heard some fraction of the presented signals which the cautious observer did not report. The simplest assumption is that the uncautious observer will report the same fraction of these relatively weak signals as he will of nonexistent signals, that is, that he will also claim to hear q per cent of the presented signals not reported by the cautious observer. Since the fraction of presented signals not reported by the cautious observer is $1-(p_c/100)$, we have the simple equation

$$p_u = p_c + q\left(1 - \frac{p_c}{100}\right)$$

which may be transformed to yield

$$p_c = \frac{p_u - q}{1 - (q/100)} \,.$$

This equation has been used here to give corrected values of the recognition probability corresponding to what a hypothetical cautious observer would obtain. It may be noted that when results are based exclusively on guessing, the observed percentage recognition probability p_u becomes equal to q, and the corrected recognition probability p_c vanishes. By use of this formula, the lower set of curves has been drawn in Figure 2 to give the corrected values of the recognition probability. In Figures 3 and 6 only corrected curves are given, with the observed values of p_u omitted. It will be noted that, as p_u approaches 100 per cent, p_c approaches p_u; this is partly the result of decreasing q with increasing signal-to-reverberation ratio and partly the diminished importance of q in correcting p_u as p_u increases.

The values of n for the corrected transition curves shown by the lower lines in Figure 2 is about 1.7. This is somewhat smaller than is typical for tests in which guessing is penalized (see Figure 10). The more gradual decrease of detection probability in the present tests may be due in part to the fact that a confident observer is more likely to identify correctly weak signals to which a cautious listener would not respond.

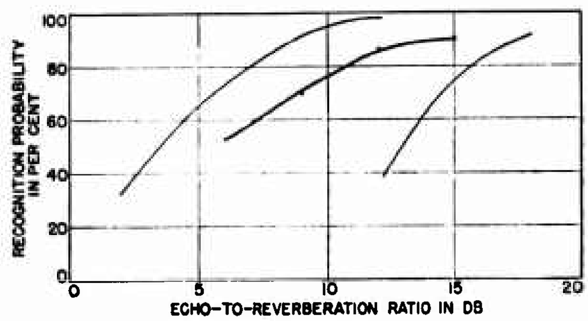


FIGURE 3. The detectability of 100-millisecond pulses masked by 25-millisecond reverberation is shown by the middle curve. The other curves have been transcribed from Figure 2 for comparison. Of these, the left hand curve refers to recognition of a 100-millisecond pulse masked by 100-millisecond reverberation, and the right hand curve to recognition of a 25-millisecond pulse masked by 25-millisecond reverberation.

The middle curve in Figure 3 shows the results obtained by presenting a 100-millisecond pulse mixed with 25-millisecond reverberation. Both sounds were passed through the filter already described. This form of test indicates one method by which it may be possible to segregate the effects due to different variables and thereby to arrive at a better understanding of the factors which determine performance. In addition, this test resembles the type of situation which may occur in practice when an echo is prolonged as a result of reflection or multiple-path transmission. It would also be interesting to determine the audibility of a pulse shorter than that used in producing the reverberation.

The audibility of a 100-millisecond pulse in the presence of 25-millisecond reverberation was about 6 decibels better at all signal-to-reverberation ratios than for a 25-millisecond pulse masked by 25-millisecond reverberation. Some such improvement would be anticipated from the increase of pulse length. In fact, it might be expected that, for a 100-millisecond pulse, performance would be better against a background of 25-millisecond reverberation than against 100-millisecond reverberation, since the difference in spectra, as well as the relative duration, of the rounded 100-millisecond pulse and the 25-millisecond reverberation blobs should facilitate recognition of the pulse. Actually, the observed results are some 3 decibels worse at all signal-to-reverberation ratios than for a 100-millisecond pulse masked by 100-millisecond reverberation. Without further experimental work, it is impossible to say whether this effect is real and associated with differences in reverberation structure, or the result of some systematic error, perhaps associated with the effect of the band-pass filter.

9.1.3 Recognition Differentials for Recorded Echoes

To test the practical reliability of the results obtained with rounded pulses, a study was made of the audibility of recorded echoes. These were obtained by echo ranging on an artificial target with a pulse of 24-kilocycle sound 100 milliseconds long. The echoes and accompanying reverberations were heterodyned to 800 cycles and recorded. The target used for this purpose was a triplane, consisting of three mutually perpendicular planes. This object has a very high target strength for its size, because a beam of parallel rays tends to remain parallel after reflection from it. Unlike a submerged submarine, the triplane can readily be kept at a fixed distance from the echo ranging vessel and is therefore much more convenient to use in testing work.

The echo from a triplane is usually a fairly accurate reproduction of the outgoing pulse, in contrast to echoes returned by large targets, since such echoes tend to be considerably prolonged and frequently show marked amplitude modulation not present in the outgoing pulse. Thus, these tests do not represent what would be obtained with actual echoes from submarine targets. The purpose of the tests was primarily to provide recognition differentials for pulses and reverberation produced with exactly the same gear at a single time, and thus to eliminate the possible errors introduced by use of reverberation produced with one source and pulses produced with another.

In obtaining the echoes, the triplane was placed at moderate ranges, where the reverberation was weak. The echoes stood out strongly above the accompanying background. Subsequently, these echoes were recorded on another disk with the gain varied in such a way that a series of echoes of nearly equal amplitudes was obtained.

In cutting the records which were to contain only the echoes, a gate circuit was used to admit the echo. This gate was tripped by the beginning of each reverberation. Since the time interval between the beginning of the reverberation and the occurrence of the echo varied somewhat over the course of the original record, it was necessary to adjust the "open" time of the gate to a value considerably longer than that of the echo in order to transmit each echo completely. Hence, all recorded echoes were preceded and followed by short intervals of reverberation. Figure 4 shows the envelopes of these reverberation remnants when the rerecording was made using a "square" gate, and Figure 5 shows the results obtained with a "round" gate. It will be noted from these figures that the envelopes of the echoes are changed somewhat by either method of rerecording.

The reverberation disk was also rerecorded from the original, in such a manner that it contained only reverberation and was devoid of echo. This was done by blanking out that portion of the original trace which contained the echo and also all subsequent reverberation. This change was of no consequence, since in the tests the echo was injected at a considerably shorter range than that at which it occurred in the original record. Because the weak terminal portion of the reverberation samples was blanked out, it was unnecessary to use the 3,500-cycle low-pass filter to eliminate the effects of record noise over the weak portions of the reverberation.

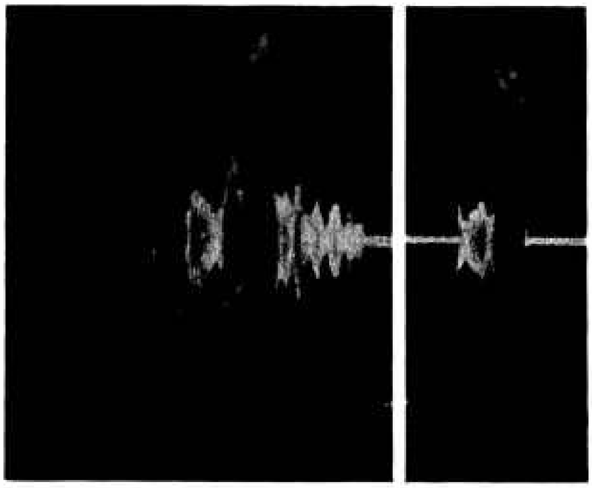


FIGURE 4. At the left is a CRO trace of an echo-reverberation mixture. To the right is shown the square-gate "echo" used to produce the mixture.

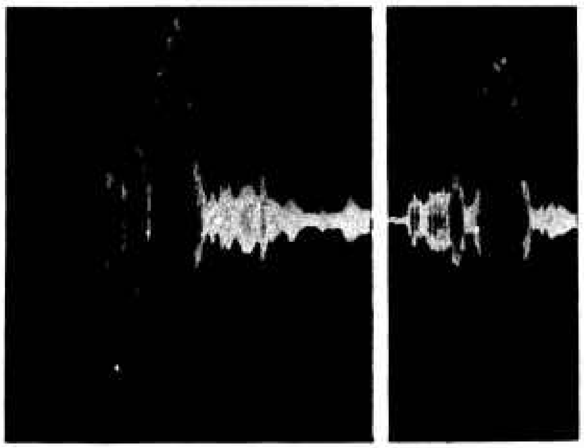


FIGURE 5. At the left is a CRO trace of an echo-reverberation mixture. To the right is shown the round-gate "echo" used to produce the mixture.

To reproduce the echo and reverberation during the masking tests, the echo disk was played in synchronism with the mated reverberation disk from which the echoes had been taken. The time interval between successive echoes in such a series was adjusted in such a manner that each echo would then coincide with the particular reverberation sample from which it had been obtained, although at a shorter range. This procedure was adopted in order to minimize spurious deviations from the practical situation, such as are associated with slow drifts of heterodyne frequency and other types of unstability which might supply false cues in the listening tests. Since the mated disks were cut from one original and were played at the same speed, the range relation between echoes and reverberations could be fixed at the outset by means of an angular scale and did not vary appreciably during a given playback.

The first result of this program of measurements was the marked dependence of recognition differential on the type of gating circuit used. The echoes which had been taken off the original record by use of the square gate were audible at a level 7 decibels lower than the corresponding pulses rerecorded by use of the round gate. This finding is apparently the result of the steep wave front transients produced when the square gate was used. These transients, which resulted from the weak reverberation before and after the echo, gave audible components at frequencies where little masking background was present. As expected, it was found that passing both the signal and the reverberation through a 140-cycle filter eliminated this spurious cue; thus, when this filter was used, the audibility of the transmitted echoes was very nearly equal to that of the rounded pulses (see Figure 2).

The detailed results obtained with the round-gate echoes and with the square-gate echoes plus filter are summarized in Figure 6. In this figure the echo recognition probability in per cent is plotted against the echo-to-reverberation ratio in decibels. The continuous line in this curve represents results obtained for injected pulses of 100 milliseconds duration, discussed in the previous section (see Figure 2). It is evident from this figure that the points obtained when the round-gate echoes, with or

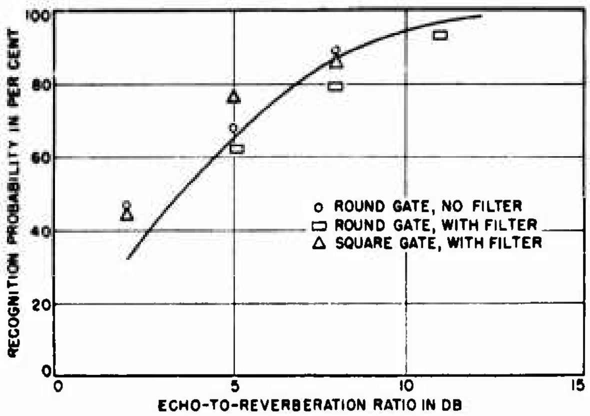


FIGURE 6. Effect of the gate and filter on the detectability of recorded echoes. The curve is transcribed from Figure 2 and is seen to fit the points moderately well.

without filter, were used agree well with the corresponding results found with the square-gate echoes, plus filter, and that both of these agree with the data for artificial pulses.

9.2 UCDWR TESTS

To obtain detailed information on the masking of echoes by reverberation under a wide variety of situations, an extensive program was undertaken during the war by UCDWR. The objective of this program was primarily to determine recognition differentials of practical importance for echoes masked by reverberation under controlled and varied conditions, and secondly to cast light on the basic factors affecting echo masking. The chief variable of interest was the pulse length, since this quantity can readily be varied in the field to give optimum results. Other variables investigated were the range at which the echo appeared, the loudness level, and the amount of distortion introduced. The general procedure followed in this program is described in Section 9.2.1, while the results obtained are given in Section 9.2.2.[b]

[b] All the descriptive material in this section has been informally communicated. A report on these tests is expected to appear in the near future.

9.2.1 General Procedure

These masking tests were made with recorded beam-aspect echoes (no doppler) and recorded CW reverberations produced by pings of the same duration as that of the echoes to be detected. The 50 per cent recognition differential for each test was found by plotting a transition curve giving percentage recognition obtained by a group of five observers for each echo-to-reverberation ratio.

ECHO AND REVERBERATION RECORDINGS

The echoes were obtained from an S-class submarine at close range. The sound-in-the-water, for reverberations as well as echoes, had a frequency of 24 kilocycles and was heterodyned to 800 cycles before being recorded on film. Recorded echoes were selected for this study which were fairly clean and which stood out well above the accompanying reverberation. Sample echo recordings are shown in Figure 9 of Chapter 8. Furthermore, all the echoes used had essentially the same duration as the emitted pulse, as established by measurements made on the film. Transmission lengths of 11, 36, 97, 114, and 271 milliseconds were employed to obtain the echoes and reverberations studied in the present tests. Some of these same echoes were also used in the thermal noise tests described in Section 8.4.6.

The echo, together with accompanying background, was recorded on the sound track of motion picture film. The remaining portion of the film, normally reserved for photographs, remained blank and was used to control the gating circuit which admitted the echo to the signal channel. By applying opaque tape to the blank portion of the film, except between the points corresponding to the beginning and end of an echo, and designing a photoelectric circuit to actuate the signal channel only during the course of the echo proper, it was found possible to eliminate spurious cues due to transients. Clearly, such recorded echoes are not entirely free from noise and reverberation background, but it is probable that contributions from these sources had a negligible effect on the results.

In order to minimize the effects of noise inherent in the recording processes, recordings were made at a fairly high gain setting; it is possible that some rounding of the echo and reverberation envelopes occurred whenever the received sound level exceeded the dynamic range of the recording equipment. Annoyance and fatigue due to film noise were virtually eliminated by protecting the recording from dust and scratches. As a further precaution, the echoes and reverberations were passed through a common band-pass filter, with cut-offs at 565 and 1,130 cycles, respectively. This filter was wide enough to transmit the sounds without substantial distortion of their envelopes. Tests with and without the filter gave the same recognition differentials for given signal-background combinations.

The sensed quality of the longer echoes was markedly tonal, and they had a definite and steady pitch. The shorter echoes were less tonal and more "noiselike"; thus, echoes with durations of 10 milliseconds or less tended to sound like a crack or pop.

Recordings were made of fairly long deep-sea reverberations. The dynamic range of the receiving system and the film, and the low level of inherent noise which can be achieved, permitted high-fidelity recording of reverberation background between ranges of about 100 yards and ranges exceeding 2,500 yards. Since the reverberations were recorded without any time variation of gain, the lower limit of 100 yards was imposed by overloading of the film for this and shorter ranges. This is a realistic facsimile of conditions encountered when using practical receivers which are not equipped with level stabilizers, such as *reverberation-controlled gain* [RCG] or *time-varied gain* [TVG]. The upper limit of about 2,500 yards was imposed by the level of prevailing deep-sea ambient. In the tests described in this section, the gain in the background channel was held fixed during the course of each reverberation so that the average level of background was initially fairly high and gradually diminished. Figure 7 shows oscillograms for two reverberations produced by CW transmissions of the indicated duration. Owing to the expanded time scale used in preparing the oscillograms, only short sections of these reverberations are reproduced. In addition, it should be noted that the oscilloscope shows reverberation amplitude on a linear deflection scale, whereas the power level traces (Figure 5 in Chapter 7) show reverberation intensity on a decibel scale.

11 MILLISECONDS

114 MILLISECONDS

FIGURE 7. Much-reduced CRO traces of CW reverberation over a 300-yard range interval (0.38 second). The horizontal timing marks, visible below the reverberation traces, were used to determine the ranges at which echoes were injected and the corresponding reverberation levels.

Between successive reverberation recordings, the output frequency of the receiver was varied in 20-cycle steps, and a great many samples were recorded, in order to provide an adequate supply of backgrounds which would match the pitch of the selected echoes and which would show no frequency drift during their courses. The salient pitch of many samples of recorded background apparently changed progressively during the course of the reverberation decay. Such samples were not used in the masking tests. Measurements on the film and aural matching of the pitch of echoes and reverberations to that of an oscillator tone of known frequency showed that in no case did the frequencies of the echoes and the reverberations which were finally selected for test purposes differ from each other by more than 5 cycles.

Reverberation records were also selected for absence of marked echo-like blobs as well as transmissions from other echo-ranging vessels. It was found that such "clean" reverberations gave more consistent results, with fewer false echoes reported. This careful selection of reverberation samples is obviously not matched under field conditions. The results consequently apply to the ability of an operator to hear echoes in the presence of masking reverberation. The problem of distinguishing between echoes from submarine and nonsubmarine targets is a separate one and has received separate study.[5]

The reverberation playbacks, for the longer transmissions, were ringing and tonal in character, and fluctuated markedly in loudness during the period of their general decay. Frequently, these fluctuations, or blobs, sounded very much like echoes, and in preliminary tests were commonly mistaken for echoes. For shorter transmissions, the ringing character became progressively less noticeable and the sound acquired a noisier quality. The loss of tonality is greatest for reverberations produced by pings less than 10 milliseconds in duration, and the sound in this case more nearly resembled that of a reverberant gunshot. A narrow-band analysis made of a 100-millisecond CW reverberation with the aid of a 5-cycle filter indicated a total frequency spread of about 20 cycles, in agreement with the $2/\tau$ criterion discussed in Chapter 7.

Fluctuations of intensity are still perceptible in reverberations from short transmissions but are less striking because of the small duration of the blobs and the consequent fact that the ear tends to integrate over a greater number of the shorter blobs. In other words, the perceived background is smoothed to some extent because of the ear's build-up time. In the reverberation samples obtained with transmissions 114 and 271 milliseconds long, there seemed to be less correlation between the lengths of the blobs and the lengths of the emitted pulses than previous studies[2] had indicated for the shorter transmissions.

In all but a few of the tests the sounds were presented through high-quality headphones worn by the observers, who could locally vary the total gain over a range of about 15 decibels to obtain the level of greatest comfort. The initial blast of the reverberation set an upper limit to the value of the gain which could be used for long listening periods; and, in the fixed-range tests described below, the echo was injected at a point in the reverberation decay characteristic corresponding to a loudness level of approximately 50 phons. In general, the level of the reverberation background at the instant that the echo was injected produced at least 20 decibels more masking than did room noise. Whenever excessive amounts of room noise occurred during a test, the results were rejected and a retest given. Except in the tests in which the effect of distortion was studied, the total amount of nonlinearity in the recording and reproducing systems did not exceed 1 to 3 per cent. This precaution was necessary in order to avoid introduction of spurious pitch differences.

In order to reproduce the echoes and reverberations with the same pitch, the motions of the film loops in the signal and background channels were synchronized. This procedure is important for the additional reason that it permits injection of the echo at the same range in each presentation and thereby simplifies the task of correlating echo detectability with the time-amplitude pattern of the reverberation background. Despite all efforts to obtain absolute synchronism, very slight short-term fluctuations of film speed remained, so that there were residual variations in the phase relations between echo and reverberation in successive presentations of a given echo against a given reverberation at a "fixed" range. In Figures 8 and 9, for example, are reproduced oscillograms showing the time-amplitude patterns of successive mixtures of a given echo without doppler and a given reverberation at a fixed range, and also traces of the rectified, smoothed signal obtained from the echo-reverberation mixture. The only difference between successive traces is the random variation in the level of the echo-reverberation mixture. In the first of these figures, the echo-to-reverberation ratio exceeds the value corresponding to primaudibility, and the two sounds are of different amplitudes. The total variation in level of the mixture amounted to about 2 decibels in this case. In Figure 9, the echo-to-reverberation ratio was very nearly that corresponding to primaudibility. Since, in this case, echo and

FIGURE 8. Unrectified CRO traces of echo-reverberation mixtures for a high signal-to-background ratio. The arrows point to the echoes.

reverberation were of nearly equal amplitudes, the phase interference effect was large, amounting to about 7.6 decibels, which appears to be nearly the maximum range of variation to be expected from this effect. The amount of the displacement of the two film loops relative to each other as they pass the light gate, which is required to produce a phase shift of 180 degrees, is only 0.01 inch for a sound frequency of 800 cycles. This appears to be the irreducible minimum in the variation from absolute synchronism which is inherent in the equipment used.

The major result of this phase interference effect is apparently to change the level and envelope of the presented signal, although some distortion of the effective echo envelope may also be produced. In practice, such interference effects will occur before rather than after heterodyning, but it does not seem likely that reversal of the order of these operations will significantly affect the results. As will be clear from the results reported in the following section, echo and reverberation are not likely to be of approximately equal amplitudes at primaudibility except in the case of the longest echoes. This factor may have contributed somewhat to the scatter among observed recognition differentials for the case of the 271-millisecond signal (see Figure 12).

Measurement of Signal and Reverberation Levels

The levels of echoes and reverberations were found by measuring the deflections which these sounds produced on the screen of a cathode-ray oscilloscope. In practice, the screen was photographed on a continuously moving film. Since the oscilloscope had a linear scale, the measured deflection on this special film was then proportional to the echo or reverberation amplitude at any time. This method was found preferable to making measurements on the original sound track, which was too unwieldy and which had to be protected from excessive handling. Oscilloscope records made in this way are reproduced in Figure 7 in this chapter and also in Figure 9 of Chapter 8.

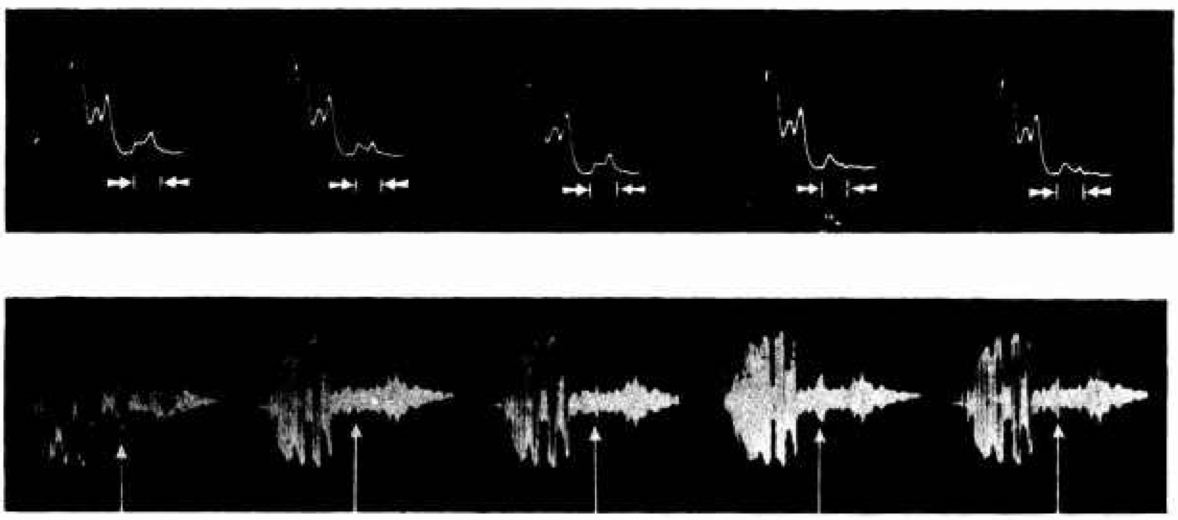


FIGURE 9. Unrectified and rectified CRO traces of echo-reverberation mixtures for primaudible signal-to-background ratio. The arrows point to the echoes. There is no correspondence between traces in the upper and lower sets.

Records such as these were used for measuring the duration and amplitude of the different echoes used. Most of the echo amplitude measurements were made at the peak levels of the echoes. For the single-point measurements (see Section 9.2.2), the levels at echo midpoint were used in computing recognition differentials. It was found for these beam-aspect echoes that the average echo level never differed from the midpoint level by more than 1 to 2 decibels.

In measuring the reverberation amplitude by this method, a time scale was required to select the proper point for measurement. This was provided by means of a 100-cycle tone which actuated at 10-millisecond intervals a neon lamp mounted at the edge of the screen. By photographing the screen and the successive flashes of the neon lamp, a film record was obtained of the instantaneous reverberation amplitude together with the time after reverberation onset, or range (see Figure 7).

Test Procedures

In any test, the echo level was varied randomly over a range of 14 decibels containing 7 steps of 2 decibels, and about 1/3 of the presentations were blanks. The onsets of successive reverberations were spaced 3 seconds apart to simulate a 3-second keying interval, which was found to allow the observers sufficient time to record their judgments of echo audibility. In this manner, signal-to-reverberation ratios could be varied from nearly certain to vanishing detection probability. Since it was found that some sequences of signal-to-reverberation ratios tend to produce inversions of performance, two successive echo levels were never permitted to differ by more than 6 decibels. For example, weak echoes are often missed when they are presented immediately following very strong ones, although subsequent weak echoes of the same magnitude as the first would generally be reported as audible. Similarly, when an audible but weak echo follows a long succession of no-echo intervals, there is a tendency to miss such a weak echo. On the other hand, a blank following a succession of strong echoes is not usually likely to lead to a commissive error unless the reverberation presented during the blank interval contains an echolike blob. Although it is important to be alert to the influence of effects such as these in fundamental laboratory studies, in order to improve the stability of the test results, it is clear that such favorable conditions cannot generally be expected in the field. Hence, it may be valuable in further studies to obtain quantitative estimates of the probable magnitudes of such effects in determining detection probability under service conditions.

Ordinarily the total duration of each test was about 20 minutes. This length of time is comparable to a half-hour watch and was found to be short enough so that fatigue effects were unimportant. While no controlled study was made of the influence of fatigue, it was found in a number of tests which were longer than usual that fatigue may have a large and unfavorable influence on the stability and excellence of performance.

It was pointed out earlier that the reverberations used in these tests were carefully selected in order to eliminate spurious effects, such as differences in pitch between echo and reverberation, and progressive drifts of reverberation frequency with time. Hence, preliminary tests were conducted in order to determine the minimum number of reverberations which could be used (for testing with an echo of given duration) without introducing significant changes in performance produced by memorizing the reverberation samples. It was found that, in general, 5 reverberation samples was a sufficiently large number so that memory effects became relatively unimportant. Thus, the *RD* observed for a particular echo, injected at fixed range, when the reverberation film loop contained 11 reverberations was only 1 decibel worse than for the same echo injected at the same range when the reverberation film loop contained only 5 of the 11 samples previously tested. When the background loop contained only 1 reverberation sample, repeated over and over again, the improvement in *RD* as compared with an 11-sample loop amounted to between 1 and 4 decibels, depending on the pulse length. Memorizing was easier for the longer transmissions because the blobs in the resultant reverberation were longer and therefore fewer in number.

Except when otherwise stated, all UCDWR tests were made with a background loop containing 5 reverberation samples. No explicit correction, which would amount to approximately 1 decibel, has been applied to these data in order to allow for the effects of memorizing the background samples, and this should be borne in mind when assessing the test results. During these tests on memory effects, it was found that often a very weak echo could be detected because its injection produced a slight but detectable change in the time-amplitude pattern of the reverberation. None of the results quoted here reflect the influence of this factor since it is, in general, quite unrealistic and all data influenced by it were eliminated from the final results. However, bottom reverberation received in anchored echo-ranging gear, such as is used in the protection of harbor entrances, may have a structure which tends to recur in successive reverberations. In such cases, recollection of reverberation envelope might play a significant role in submarine detection.

The echoes used in most of these tests were injected at fixed range in order to simplify the study of correlation between audibility and reverberation envelope; it was this aspect of test procedure which necessitated caution in minimizing the effects of memory. One listener who participated in a small number of tests had a highly developed sense of rhythm and could tell from the rate of succession of the blobs the precise instant at which to concentrate on the echo in the case of the fixed range tests. However, as pointed out above, there were usually so many variations in reverberation envelope and they followed each other so rapidly that mere knowledge of the approximate range was not a very helpful cue to the average observer when the number of reverberation samples exceeded 5. This is confirmed by the fact that recognition differentials obtained in tests in which the range was varied are in good agreement with the fixed range recognition differentials.

The average incidence of commissive errors occurring in all tests in this program was 3.5 per cent. The greatest incidence was noticed in the 11-millisecond tests, where the relative number of such errors rose to 7 per cent. However, even for the 11-millisecond echo no more than 1.5 per cent of commissive errors were made when a different and "cleaner" set of reverberations was substituted. The tendency toward commissive errors was minimized by interviewing and retesting observers whose votes showed a bias toward multiple commissive errors.

Training Effects

The observers used in these tests were laboratory personnel, many of whom had previously participated in tests in which the masking of target sounds and of recorded echoes by noise background were studied. Generally, the tests were administered to groups of five observers. In standard psychological tests of auditory acuity, the Seashore tests, the group scores were found to fall in the highest decile for pitch discrimination and in the second decile for loudness discrimination. One member of the test group possessed an accurate sense of absolute pitch. Examination of absolute auditory thresholds showed no significant hearing loss up to 10 kilocycles for any member of the test group. Comparison of recognition differentials obtained in one series of reverberation masking tests, in which the regular test group and also five sonar instructors attached to the West Coast Sound School participated, showed that the performance of both groups of listeners was the same to within 1 decibel, and these small differences showed no systematic trend within the test series on which the comparison was based.

It should be pointed out, however, that in effect both of these groups consisted of experienced listeners. The UCDWR observers were given a thorough preliminary orientation in the nature of the sounds to be used in the tests, since it was noticed early in the test program that failure to supply such preliminary experience often resulted in failure to report fairly strong echoes, especially for short ranges and pulse lengths, or produced consistent errors of commission. This phase of listening indoctrination produces an understanding of the problem. It involves a fairly abrupt change in listening performance, and once understanding exists, performance is much more stable. This kind of indoctrination is obviously important in training sonar operators.

During the course of the listening tests, the performance of observers also showed a tendency to improve slowly and progressively with increasing experience. This effect is rather important in listening studies involving elusive and subtle distinctions between echoes and blobs. Thus, when the group of listeners was shifted from tests in which beam-aspect echoes were reverberation-masked to tests in which similar echoes were noise-masked, and then after a period of a week or two were returned to reverberation tests, it was found that their performance in the reverberation tests had deteriorated by about 3 decibels from the previous standard. However, after a week or more of reverberation testing, the group performance in the reverberation tests had returned to its earlier value. This observation implies that periodic field exercises, refresher courses, or the opportunity to listen to recordings of typical echo-reverberation mixtures would be valuable in maintaining a high standard of performance among sonar operators.

Since all the results reported were obtained under conditions in which the effect of experience was relatively constant, comparisons between them do not reflect the influence of this factor to any significant extent. However, it should be pointed out that there is an irreducible minimum of variability in performance arising from the nature of the material selected for test purposes. Thus, it was observed in these tests that two different echoes of the same length and general level, but with somewhat different envelope shapes, were not equally audible in the presence of reverberation. Similarly, the variability in performance for this group of listeners was considerably less in the reverberation masking tests than in the noise masking tests reported in Section 8.4. This is presumably due in part to the fact that the noise tests were preliminary in nature and the total testing period too brief to permit the listeners to acquire enough experience to give very stable performance.

It is interesting to note in this connection that one of the effects of training is to make the responses more nearly automatic. Thus, it was found that experienced observers usually would do just as well regardless of whether or not they made an active effort to concentrate on each item; in other words, rather passive listening could apparently be tolerated without significantly affecting performance. This finding implies that the results obtained in these laboratory tests may be applicable to the field situation despite the greater number and variety of distractions encountered in practice.

9.2.2 Observed Recognition Differentials

Fixed Range Tests

Recognition probability curves for the five echo lengths studied in these tests are shown in Figure 10. In this case, the echo-to-reverberation ratio shown on the horizontal scale is

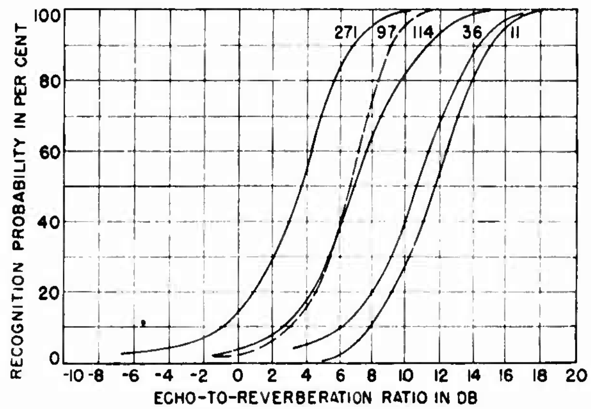


FIGURE 10. Probability of detecting an echo masked by reverberation. Each curve is an average of five transition curves obtained with an echo of the indicated duration in milliseconds.

defined in terms of reverberation and echo levels measured at the midpoint of the echo. These curves are composites, showing the mean performance of groups of five observers in tests in which the echo was injected at three different ranges varying from about 900 to about 2,000 yards. For each of these three ranges, 15 separate determinations of recognition probability were made. For each pulse length, however, only a single echo was used. Failure of these curves to approach zero as rapidly as they approach 100 per cent recognition probability is due to the influence of commissive errors, although as shown by the curves, and as pointed out above, the effects of guessing were relatively small in these tests. Transition curves given by observers who try to force their thresholds down tend to be much more gradual in slope than are the curves shown in Figure 10, and such transition curves often show inversion at the lower signal-to-background ratios (see Figure 7D in Chapter 4). The mean value of n for these transition curves is about 1.8.

The data on which the transition curves shown in Figure 10 are based have been plotted in a slightly different way in Figure 11 in order to show the fractional part of the test group for which the echo was audible in at

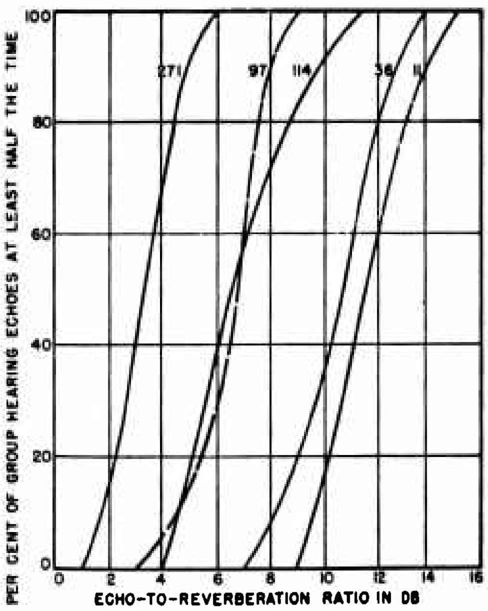


FIGURE 11. Fraction of test group detecting echoes in at least 50 per cent of trials for echoes of the indicated durations in milliseconds.

least 50 per cent of the trials at various values of the signal-to-reverberation ratio. It will be seen that the difference in performance of the best and poorest observers did not exceed 6 to 7 decibels and that this difference did not depend strongly upon pulse length. Nearly identical results were obtained when the quantity plotted represented the fraction of the test group which could detect at least 20 per cent and also at least 80 per cent of the echoes at various signal-to-reverberation ratios.

The 50 per cent recognition differentials found by averaging the individual recognition differentials from each transition curve are shown as filled-in circles in Figure 12. Each point represents an average of 45 determinations (15 tests at each of 3 ranges) made, however, with a single recorded echo. The length of the vertical lines drawn through these circles represents the rms deviation from the mean of the results which occurred in the individual tests. The numerical values of these rms deviations are listed in Table 2 and are discussed more fully at that point. These rms deviations represent the total effect due to all sources of variation.

The inversion in recognition differentials for the 97-millisecond and 114-millisecond echoes was checked and confirmed in an additional series of tests. This anomaly may be associated with the envelope of the 114-millisecond echo. In fact, for one of the 114-millisecond echoes studied in the retest series, the *RD* was considerably lower than for the 114-millisecond echo used in the main body of this work, although the same set of reverberation samples was used in both cases. It may also be noted that the 11- and 97-millisecond echoes and reverberations were recorded on a different day from that of the other sounds. Consequently sea conditions may have affected reverberation character for the 11- and 97-millisecond tests, lowering the *RD* and increasing the rms deviation for both of these cases with respect to the other three shown in Figure 12.

The continuous line in Figure 12 has been drawn as a curve of best fit for the filled-in

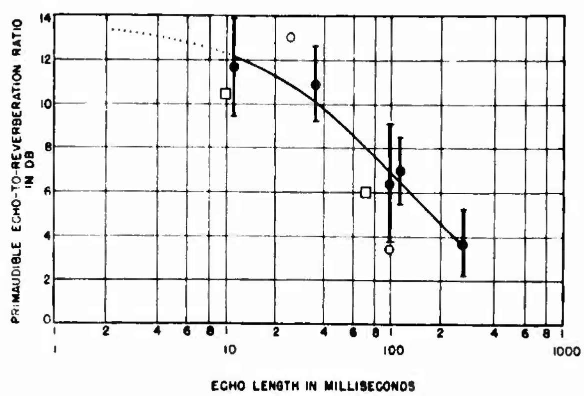


FIGURE 12. Aural detection of CW pulses and echoes at 800 to 1,000 cycles, masked by reverberation from transmitted CW pulses equal in length to the signal length and heterodyned to the same audio frequency. The filled-in circles apply to 800-cycle beam echoes, for which the recognition differential is defined by measuring the echo and reverberation at echo midpoint; the vertical lines represent standard deviations among the recognition differentials for specific echoes. The open circles apply to filtered 800-cycle pulses, for which the recognition differential is defined by measuring the average level of a group of reverberation samples at echo midpoint by means of a power level recorder. The open squares apply to 1-kilocycle injected pulses, for which the recognition differential is defined by measuring the levels of neighboring reverberation peaks.

circles representing the UCDWR results. The dotted portion of this curve has been drawn on the basis of some very preliminary data obtained at UCDWR with signals less than 10 milliseconds long. These preliminary tests indicate that the audibility of short signals masked by reverberation deteriorates much less rapidly with diminishing pulse length than does the audibility of longer echoes.

For comparison with the continuous line in Figure 12 the open symbols represent recognition differentials obtained in other programs. In particular, the open circles represent the results obtained by BTL described in Section 9.1.2, while the open squares give the recognition differentials obtained by the British (see Section 10.2.3). The agreement between these four points and the curve based on the UCDWR data is not very impressive, but all the points lie within 3 decibels of the curve shown in the figure. The British data exhibit the same change with pulse length as do the UCDWR results. Since the absolute values of the recognition differentials obtained by the British are somewhat uncertain, this agreement is as close as could be expected. The BTL results, on the other hand, show a change of 9 decibels as the pulse length decreases from 100 to 25 milliseconds, and this change is much greater than can be reconciled with the UCDWR results. Since the latter extend over a greater range of pulse length, they are believed to be more reliable. The discrepancy is puzzling, however, and casts some uncertainty on the reliability of the curve of best fit shown in Figure 12. Possibly the fact that the 140-cycle filter used in the BTL tests was centered 35 cycles away from the midfrequency of pulse and reverberation may have impaired recognition of the 25-millisecond pulse relative to the 100-millisecond pulse.

The continuous curve in Figure 12 may be used to estimate the advantage to be expected by reducing the pulse length when reverberation is the dominant element of background. As stated in Section 7.3, the average intensity of reverberation background at a given range (for fixed oceanographic conditions and power output) is inversely proportional to pulse duration. In other words, reverberation intensity is proportional to the energy in the emitted pulse; hence, a tenfold reduction of pulse length results in a 10-decibel drop of reverberation level for a specific set of conditions. Thus, the fact that the recognition differentials shown in Figure 12 deteriorate at a rate less than 10 decibels per tenfold decrease of pulse length means that weaker echoes can be detected in the presence of reverberation, everything else being the same, when the pulse length is shortened. Figure 13 gives an estimate of the

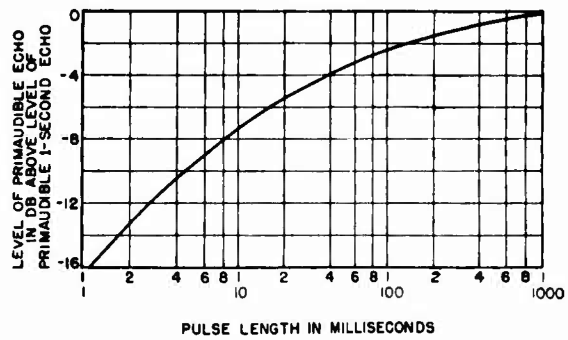


FIGURE 13. Estimated improvement in echo detectability resulting from decreased pulse length, when reverberation is limiting.

amount of this improvement and is based on the probable recognition differentials shown by the continuous line in Figure 12. According to Figure 13 there is an estimated improvement of about 5 decibels in changing from a 100-millisecond to a 10-millisecond transmission. This estimate is in agreement with the practical observation[3] that shortening the transmission is helpful in the presence of reverberation. However, loss of ability to distinguish doppler and thus to discriminate wake from target echoes makes a short transmission undesirable for search when the echo is to be detected by ear, although shortening the pulse may be helpful during action and is probably somewhat more helpful when nonauditory detectors are used. It should be noted that these results are valid only for clean echoes and are thus applicable only to submarines at beam aspect. At other aspects, smear echoes appear, the echo becomes prolonged, and in addition its mean intensity decreases proportionally with the pulse length. The relative advantage for short pulses shown in Figure 13 is, therefore, not generally applicable.

It is interesting to compare the recognition differentials found for reverberation-masked echoes with those given in Chapter 8 for noise-masked pulses. This comparison may be expected to cast light on the fundamental processes involved in auditory detection of pulses. One important difference must be noted before this comparison can be made. When a pulse is masked by reverberation, the measured signal-to-reverberation ratio gives directly the ratio of signal to reverberation in each of the ear's critical bands, which are stimulated, since the spectrum of signal and reverberation are almost identical, unless own-doppler broadens the reverberation spectrum by a large amount (see Section 10.1.2). With a noise background, on the other hand, the spectra of signal and background are quite dissimilar, and an assumption must be made to determine the signal-to-background ratio in a critical band. Specifically, the equivalent filter width of a critical band must be known so that the intensities of signal and background in this band can be computed. In Section 8.4.4 the signal-to-background ratio of the primaudible pulse in one of the ear's critical bands has been computed and the result plotted in Figure 11 of Chapter 8.

This computed curve showing the ratio of primaudible signal to noise in one of the ear's critical bands has been transcribed to Figure 14 where it is shown as the upper dashed

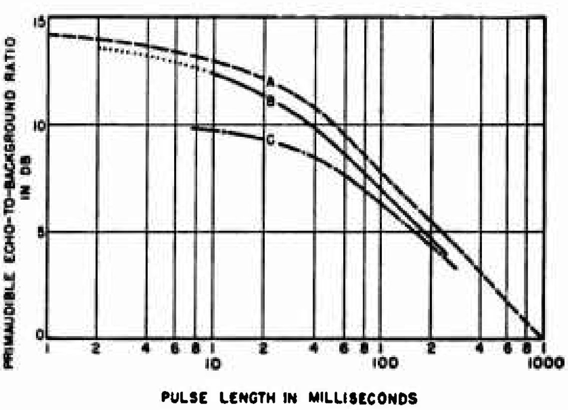


FIGURE 14. Comparison of recognition differentials for 800-cycle signals masked by noise and reverberation. Curve *A* shows critical-band recognition differentials for wide-band noise backgrounds (Figure 11 in Chapter 8). Curve *B* shows recognition differentials for reverberation measured by the single-point method (Figure 12), while curve *C* shows recognition differentials for reverberation measured with an electrical time constant of 100 milliseconds (see Table 3).

curve. The continuous curve for reverberation masking is also given, transcribed from Figure 12. Since the two curves were prepared quite independently of each other, the parallelism shown is striking. This figure gives added support to the inference, stated in Section 8.5.3, that the critical-band *RD* for a very brief pulse tends to be independent of pulse length, approaching a constant value of about 14 decibels. This result is rather surprising and at present must be regarded as suggestive rather than conclusive. In particular, more data on reverberation masking of pulses about 1 millisecond long are required to test the reality of this apparent result.

Curve *C* in Figure 14 represents the recognition differentials found by electrical integration of the reverberation level, which are discussed below. The failure of this curve to agree with the curve *B* casts some doubt on the validity of either and suggests that the tentative agreement between the curves of *RD* for reverberation and noise masking may not be real. However, as pointed out below, it is difficult to understand why the electrical integration gives such systematically different results from those obtained by the single-point method. Evidently more data are required before definite conclusions can be drawn.

The fact that the reverberation recognition differentials shown in Figure 14 are somewhat smaller than the critical-band recognition differentials obtained in the presence of wide-band noise may be accidental. If corrections were introduced for the effects of memory and for use of a decibel, or power level, rather than a power average in obtaining mean recognition differentials for reverberation masking, the latter values (at least those taken from Figure 10) would probably tend to lie somewhat above the critical-band recognition differentials for noise masking. Also, the loudness level per critical band was considerably less when wide-band noise was used than for the reverberation tests, since the overall backgrounds probably had much the same loudness in both sets of tests. Since recognition becomes more difficult at lower loudness levels (see p. 231), this effect would account for an increase of noise-masking recognition differentials compared to those for reverberation masking. It would be expected that when all these factors were taken into account, the critical-band recognition differentials for reverberation should exceed those for noise, since misleading cues like reverberation blobs are much less likely to occur in the case of masking by wide-band noise. However, the *RD* at 10 milliseconds found by electrical integration (see Table 3) seems too far below the noise masking *RD* to be brought above it by any of these corrections.

VARIABILITY

The individual recognition differentials which were averaged to give the plotted points in Figure 12 showed a high degree of variability. The origin of this variability was the subject of considerable study at UCDWR, and the results obtained cast some light on the basic mechanism producing echo masking.

The nature of this variability was found to be intimately connected with the manner in which the reverberation intensity was measured. For the data plotted in Figure 12, instantaneous levels of echo and reverberation were measured at a time corresponding to the midpoint of the echo. This method is referred to hereafter as the *single-point* method. As shown in Figure 12 and Table 2, the rms deviation among recognition differentials calculated by the single point method is of the order of 4 decibels; hence, the total range of variation is about 12 to 14 decibels, and extreme deviations from the mean may be as large as 6 to 8 decibels.

The arithmetic meaning of this scatter is best understood by referring to Figure 15. In this figure the level of the primaudible echo is plotted against the reverberation level measured by the single-point method. The different points represent measurements made with a single 79-millisecond echo against eleven different reverberations, at each of two ranges. It is clear from this figure that the reverberation level measured by the single-point method seems to have only a general relationship to the level of the echo that can just be detected.

If the primaudible echo level were equal to the reverberation level plus the recognition differential, the points in Figure 15 should lie close to a 45-degree straight line, deviations from the line resulting from inaccuracy of the measurements. Some curvature of the line would be expected if the *RD* varied with the loudness level, but for the loudnesses used this effect was probably negligible. Thus the 45-degree straight line shown in Figure 15 was drawn. It will be noted that this line intersects the vertical axis at about 10 decibels, within 2

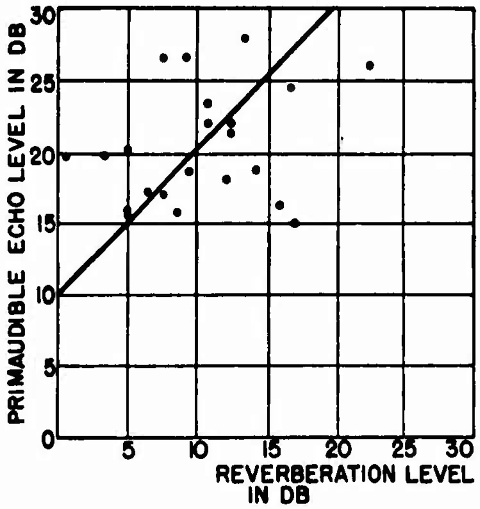


FIGURE 15. Recognition differentials computed by the single point method. The two sets of symbols apply to different echo injection ranges. Echo and reverberation levels are given in decibels above the same arbitrary reference level.

decibels of the mean curve drawn in Figure 12 from more complete data. This intercept is clearly the *RD*, since the latter is the difference between signal and background levels.

The scatter from this straight line is so large, however, that doubt is cast on the meaning of recognition differentials found with the single-point method. All the obvious possible causes of such scatter are ruled out, however. The physical measurements of echo and reverberation level, following the techniques described in Section 9.2.1, were quite objective, and were reproducible to within a small fraction of a decibel. Similarly, when different observers were used or when a test was repeated with the same observer, the level found for the primaudible echo injected at a particular range into a particular reverberation was moderately reproducible. In fact, the standard deviation of the primaudible echo determined for a fixed reverberation at a fixed range varied between 1.0 and 1.2 decibels. This average scatter of 1.1 decibels represents the variability of the ear's performance. As a result of this irreducible scatter, the recognition differential for a particular presentation was measured with a scatter of only 1 decibel, or ¼ of the standard deviation found when the echo was injected into different reverberations or at a different range in the same reverberation.

The conclusion may be drawn that the single-point method does not adequately measure the masking power of reverberation. In view of the rapid variability of reverberation and the integrating property of the ear, this is not a surprising result. We would expect that the masking power of reverberation would be determined by the average reverberation level, taken over an interval of several hundred milliseconds, that is, over an interval comparable with the build-up time of the ear. Therefore other methods were devised to measure the acoustically significant reverberation level in an attempt to reduce the variability of the measured recognition differentials.

First, the reverberation level was measured at 5 points equally spaced over the echo, at the beginning, the end, and three equally spaced points during the echo. The recognition differentials obtained by this method, together with the standard deviation of recognition differentials, is given in Table 2. For comparison, the standard deviation of reverberation intensities from the mean intensity at each range is given on the right-hand side of the table. All the corresponding data for the single-point method are also given in this table. It is evident that averaging the reverberation intensity over 5 points considerably reduces the standard deviation both of the individual reverberation levels and of the calculated recognition differentials.

This same point is also illustrated in Figure 16. Here the primaudible echo level is plotted against the reverberation level for each back-

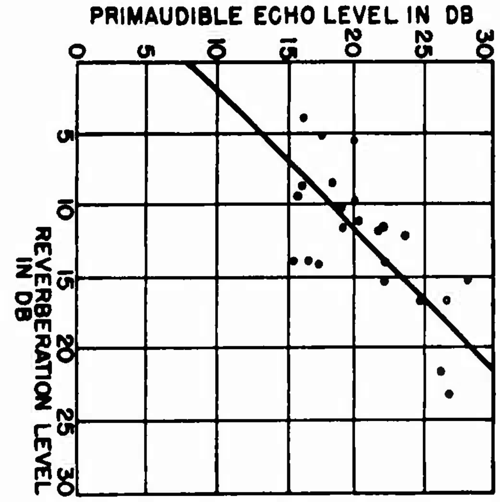


FIGURE 16. Recognition differentials computed by the 5-point method. The two sets of symbols apply to different echo injection ranges. Echo and reverberation levels are given in decibels above the same arbitrary reference level.

ground sample and at a given range with each of 11 different samples. The points correspond to the same echo levels as in Figure 15, but the reverberation levels were measured by averaging over the echo. It is evident that the average reverberation level found in this way is a better indication of the masking power of reverberation than is the level at a single point.

TABLE 2. Recognition differentials determined by point and band methods.

Echo length in milliseconds	Recognition differential in db		Standard deviations in db: Recognition differential		Standard deviations in db: Reverberation level	
	1 point	5 points	1 point	5 points	1 point	5 points
11	11.7	11.8	4.6	3.5	4.5	3.6
36	10.9	10.3	3.4	1.9	3.3	1.7
97	6.4	7.2	5.4	2.9	5.1	2.4
114	7.0	6.6	3.1	1.7	2.7	1.3
271	3.7	5.1	2.9	2.1	2.4	1.2

Various other methods of measuring the

reverberation level were tried. Five points prior to the echo were averaged; also a planimeter was used to integrate over the reverberation intensity prior to and during the echo. In addition, a power level recorder was used to smooth the reverberation level. Each of these significantly reduced the scatter of the measured recognition differentials. The greatest reduction in scatter was obtained with what is in practice a very simple method. An electrical system was used to smooth the reverberation; thus, the emergent signal at any time t_1 represented an integral over the reverberation intensity at times prior to t_1, multiplied at each time t by $e^{-(t_1-t)/t_0}$. The resulting integral was divided by t_0 to give an average reverberation intensity at the time t_1. Such average intensities were computed for various values of the time constant t_0, and the values at the echo midpoint were used in determining recognition differentials. The resultant values of the recognition differential are given in Table 3, for different values of t_0. As in Table 2, the observed standard deviations are also given both for the milliseconds prior to echo midpoint. This result is, of course, generally consistent with expectation.

The recognition differentials given in Table 3 are not directly comparable with those for the single-point method shown in Table 2 and Figure 12. Since the mean reverberation level decreases rapidly with time, the single-point recognition differentials are affected by the high reverberation prior to the echo. If the primaudible echo level is to be predicted from the mean reverberation intensity at the range of the echo, the single-point or five-point recognition differentials are probably valid, provided that the reverberation is heard directly, without any TVG, AVC, or similar device to maintain the reverberation level constant. On the other hand, if the reverberation level presented to the ear does not change systematically with time, the recognition differentials in Table 3 are probably more realistic. It will be noted that these give an even greater relative advantage for the shorter pulse lengths than that shown in Figure 12.

TABLE 3. Recognition differentials determined by electrical integration. Time constant t_0 of circuit given in milliseconds.

Echo length in milli-seconds	Recognition differentials in db				Standard deviations in db							
					Recognition differentials				Reverberation levels			
	$t_0=200$	$t_0=100$	$t_0=50$	$t_0=5$	$t_0=200$	$t_0=100$	$t_0=50$	$t_0=5$	$t_0=200$	$t_0=100$	$t_0=50$	$t_0=5$
11	7.8	9.7	10.1	10.5	1.5	1.4	1.3	2.7	0.7	0.9	1.2	2.7
36	6.7	9.1	9.9	10.0	0.9	1.1	1.4	1.6	0.6	1.0	1.5	1.5
97	3.8	5.9	7.0	8.1	1.1	1.2	1.9	3.8	1.1	1.1	1.8	3.6
114	4.3	6.3	6.7	7.3	1.1	1.3	1.8	2.2	0.8	1.2	1.6	2.0
271	1.1	3.5	4.2	4.8	1.4	1.5	1.8	1.9	0.9	1.0	1.3	1.3

recognition differentials and for the reverberation intensities at a fixed range. It is evident that the standard deviations in this table are much reduced from those in Table 2. Especially for time constants t_0 of 100 or 200 milliseconds, these standard deviations are so close to the limiting value of 1.1 decibels set by variability in the ear's performance that no further reduction in variability along these lines seems possible. The conclusion may be drawn that the masking power of reverberation is measured primarily by an integral extending 100 to 200 milliseconds preceding the echo.

From a practical standpoint, Table 3 implies that the recognition differential for an echo in the presence of reverberation should always be expressed in terms of the average reverberation taken over an interval 100 to 200 milliseconds preceding the echo. Since this is not very convenient, the reverberation levels during the echo may be used instead, together with the recognition differentials shown in Figure 12, but with some loss of accuracy. Especially when the reverberation falls off very sharply, use of the recognition differentials found in terms of the reverberation level during the echo and shown in Figure 12 may lead to erroneous conclusions, and this possibility should be kept in mind.

It will be noted that the recognition differentials shown in Tables 2 and 3 display some puzzling discrepancies. In particular, in Table 2 the *RD* found by the five-point method for a 271-millisecond echo is 1.4 decibels higher than the corresponding *RD* found by the single-point method. Values for the other echo lengths are in good mutual agreement. On the other hand, in Table 3 the recognition differentials for time constants of 100 and 200 milliseconds show a systematically different change with echo length than do the recognition differentials in Table 2. The points for the time constant of 100 milliseconds have been represented by curve *C* in Figure 14. The curve for the 200-millisecond time constant would be parallel, but would lie about 2 decibels lower. These differences arise entirely from differences in reverberation levels when averaged in different ways. There is no known reason why averaging over a fixed time interval should give lower values for shorter pulse length reverberation than for reverberation from longer pulses. Systematic differences between the reverberations used might explain these discrepancies, although direct measurement seemed to indicate no systematic differences greater than 1 decibel between the reverberation level at echo midpoint and the level 50 milliseconds previous. Some systematic error may affect all these measurements made with the electrical filters. It may also be noted that the recognition differentials obtained by use of the filter do not agree with the critical-band recognition differentials for noise, shown by curve *C* in Figure 14. More data are required to clarify these differences between the recognition differentials determined in different ways.

Measurements of recognition differentials were also made using a power level recorder to define the reverberation level at echo midpoint. The recognition differentials found in this way are given in Table 4. Since the method of reverberation measurement used was closely similar to that employed in the BTL tests and described in Section 9.1.1, these results are included here as a check on the validity of reverberation measurements made with the power level recorder. It is evident from Table 4 that the recognition differentials found are in substantial agreement with those obtained by use of the electrical filter, with a time constant of 100 milliseconds (see Table 3 and Figure 14). Thus the high *RD* found at 25 milliseconds in the BTL tests probably cannot be explained as a result of the method of reverberation measurement used.

TABLE 4. Recognition differentials determined with high-speed power level recorder.

Echo length in milliseconds	Recognition differentials in db	Standard deviations in db	
		Recognition differential	Reverberation level
11	13.0	1.5	0.6
36	8.5	0.8	1.1
97	7.8	2.3	1.8
114	6.2	1.5	0.7
271	4.0	1.3	0.9

It should also be mentioned at this point that the slopes of the composite transition curves based on single-point measurements and shown in Figure 10 are in general very nearly the same as those of the transition curves computed by segregating the echo recognition probabilities and plotting individual transition curves for each of the reverberations on a given test loop or for different ranges of the same reverberation. In other words, the absolute values of echo-to-reverberation ratios leading to a particular detection probability depend on the method of measuring reverberation level, but the rate of change of detection probability with increasing signal level is much the same for all reverberations and all ranges.

The significance of the preceding discussion may be reviewed with the aid of the standard deviations of the measured reverberation levels also shown in Tables 2, 3, and 4. From these, it will be noted that, when the levels of the various reverberations are measured with an instrument having an integration time approximately that of the ear, the scatter among the measured levels for different reverberations at a given range is much reduced. Similarly, recognition differentials computed on the basis of integrated levels show about the minimum degree of scatter resulting from variable performance of the ear when presented with a constant physical situation. For practical purposes, the probable performance of the ear can best be predicted with the aid of an average of single-point reverberation levels measured at echo midpoint or by using integrated levels measured at echo midpoint, and therefore indicating the level a short time in advance of echo midpoint. These two sets of recognition differentials may differ by several decibels because of the gradual decay of average reverberation level with time.

Finally, it should be mentioned that the levels of successive echoes received at sea under apparently identical conditions may vary widely from the mean level of a group of such echoes. It follows that, in order to predict the detection probability for a particular echo, it is necessary to combine the echo distribution function, which describes the probability that an echo will have a given level under specified conditions, with the transition curve specifying the probability that an echo of given intensity will be heard in the presence of a particular noise or reverberation background.[4]

Initial Contact Experiments

These tests were administered just prior to the fixed-range tests described previously. The same echo and reverberation loops were used in both types of test, the major difference being that the listeners did not know the injection range in the initial contact studies. In the initial contact work, successive echoes were injected at increasingly higher levels, beginning with levels definitely below primaudibility and ending with the level designated as well above threshold. To control guessing, blank intervals were included in the series of presentations. As in other ascending series minimal increment tests, the minimal increment *RD* was found to be in general 1 to 2 decibels higher than the random-order *RD*. However, the differences between the two types of test did not exceed the standard deviation of differences in performance in successive random-order tests and cannot be regarded as highly significant. This is especially true for ranges at which the character of the reverberation is confusing because of the resemblance of blobs to echoes.

Fixed and Variable Range Tests

In the variable-range tests, a loop containing a single reverberation was used, and the echo level and injection range were varied simultaneously. Range was varied in 200- to 400-yard steps, and separate records were kept for the responses at each range. The mean recognition differentials derived were not significantly different from those obtained in the fixed-range tests. Since the influence of memory was greater in these variable-range tests, owing to the use of a single reverberation, it may again be inferred that ignorance of the injection range may hurt detection by 1 to 2 decibels. Again, it was found that the character of the reverberation was worse and the instability of performance greater at some ranges than at others.

Effects of Listening Level

The reverberation level was varied in steps of 10 decibels from a level 5 decibels above the threshold of hearing (as determined by room noise, rather than the ear's absolute acuity) to about 75 decibels above that threshold. The highest level used probably corresponded to between 90 and 95 phons. In order to avoid acoustic shock from the initial high level when the echo was presented at long injection ranges and at high listening levels, the earlier portions of the background on the single reverberation sample used were blanked out by means of opaque tape; thus, a larger range of listening levels could be tolerated. This procedure could not be employed in the variable-range tests, and, as Table 5 and Figure 17 show, use of level stabilizers may help by 2 to 3 decibels in practice by making it possible to use optimal gain for all ranges.

The loudness level study was made with a 114-millisecond echo only, and five observers were used. In this study, the echo level was increased progressively from a point well below threshold, and blank presentations were included. In other words, the curves in Figure 17 were obtained essentially by what is called the increasing level method in Section 4.1.5.

The effects of listening level are shown graphically in Figure 17. In general, 50 per cent recognition differentials improved progressively with increasing level up to a sensation level of about 50 decibels. The improvement (decrease) in recognition differentials with increasing level is listed in Table 5.

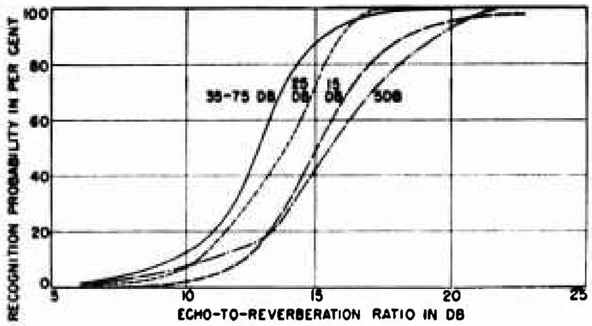


FIGURE 17. Effect of gain setting on recognition probability of a 114-millisecond echo for different sensation levels.

The performance curves for the last five levels in this table have been combined to give a single curve in Figure 17 because of the extremely small differences shown in the individual curves. The general decrease in the 50 per cent recognition differentials agrees with

TABLE 5. Effect of loudness level.

Listening level in db above lowest reverberation level audible in presence of room noise	Decrease in recognition differential in db
0	0.0
10	0.6
20	1.7
30	2.6
40	2.8
50	2.6
60	2.9
70	2.7

the trends shown in Chapter 4 (see Figures 6 through 15 and Figure 65). The rms deviation between the performance of individual observers and the mean of the group was about 2 decibels. Hence, the trend shown applies to the performance of the group, rather than to that of individual listeners. In general, commissive errors were comparatively few, but showed a tendency to increase with diminished listening level, so that the relative number of such errors was about three times larger for the lowest than for the highest listening levels.

The influence of commissive errors, together with the greater difficulty of the assigned task (recognizing a single brief echo in the presence of a rapidly fluctuating background, as compared with recognizing a periodically repeated variation in the level of a sustained sound) implies that the estimated dependence of transition curve slopes on listening level which is shown in Figure 65 of Chapter 4, is probably somewhat more reliable than the present results, at least for the case of recognition of target sounds. The total effect of loudness level on *RD* also appears smaller for the case of echoes than for that of sustained sounds and may be due to the reduced subjective loudness of brief sounds.

Perhaps the most significant inference from the loudness level tests is that RCG may prove helpful in practice. By this means a nearly optimal listening level may be maintained over the greater part of the reverberation limited range. Hence, residual masking effects due to the initially high level of received reverberation may be minimized, operator fatigue reduced, time of lost contact (due to the slowness with which manual gain adjustments are made) abbreviated, and the effects of overloading and limiting which reduce the signal-to-reverberation ratio at very short ranges, if the gain is set at an average level, virtually eliminated.

EFFECTS OF DISTORTION

The types of distortion used were the same as those described in Section 8.4.7. With moderate distortion, the mean *RD* was reduced by 0.9 decibel when the 565- to 1,130-cycle band-pass filter transmitted echo and reverberation. The standard deviation in successive tests was 1.8 decibels. When the 0.1-9 kilocycle filter was used, the mean *RD* was increased by 0.5 decibel, and the standard deviation was 0.8 decibel. Thus, moderate distortion did not introduce any very significant change in the measured recognition differentials. When the extreme type of distortion was used the *RD* was increased (caused to deteriorate) by 3.0 decibels and the standard deviation was 1.7 decibels. There can be little question that the last effect is statistically significant, in agreement with previous results. Again, the loudspeaker gave somewhat poorer results than the headphones, and the difference between the two was greatest when signal and background were admitted by the wider filter.

Chapter 10

REVERBERATION MASKING OF SIGNALS WITH DOPPLER

So far, attention has been centered upon the part played by the relative amplitudes of echo and reverberation in the process of detecting an echo without doppler. In the present chapter, the discussion is concerned primarily with the effects of pitch differences between the echo and reverberation produced by a CW transmission. Such pitch differences are produced by motion of the target and/or the echo-ranging projector through the medium and are generally termed doppler effects. Differences between the pitch of echo and reverberation which are produced by doppler are important because the masking effect of reverberation is diminished thereby. In addition, doppler shifts enable the operator to distinguish between stationary and moving reflectors and thus provide tactically valuable information.

10.1 DOPPLER EFFECT

The equations specifying the magnitude of doppler shift are well known but are derived here for reference. Certain consequences of these equations are also discussed. Consider the situation in Figure 1, which represents a

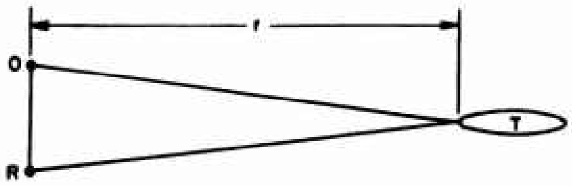


Figure 1. Reflection of sound from a target.

sustained tone of wave length λ emitted by a projector at the point O. This wave is propagated through the medium in the direction OT and is reflected back from the target T to the receiver R. If the total distance of sound travel is L, and the lateral distance between projector and receiver is much smaller than L, then L equals $2r$, where r is the range, or distance between projector and target.

When the target and the projector-receiver are stationary, r is constant, and the phase angle ϕ_R of the sound received at R differs from the instantaneous phase ϕ_0 at the projector by a fixed quantity which depends on the number of wavelengths contained in the travel path L. The difference between these phase angles is

$$\phi_R - \phi_0 = 2\pi \frac{L}{\lambda} = 4\pi \frac{r}{\lambda}. \tag{1}$$

In other words, when L is an integral multiple of λ, the receiver will lag the source by a multiple of 2π radians, and the two will move in phase.

When the source and receiver remain stationary and the target moves in a path perpendicular to the line OR, opening or closing the range at a velocity dr/dt, the difference between the phase angles at source and receiver varies with time. This rate of change may be obtained by differentiating equation (1), which gives

$$\frac{d}{dt}(\phi_R - \phi_0) = \omega_R - \omega_0 = \frac{4\pi}{\lambda}\frac{dr}{dt}, \tag{2}$$

where ω_R and ω_0 are the angular velocities of the rotating vectors representing the instantaneous phases of the received and emitted waves, respectively. In other words, ω_0 equals $2\pi f_0$ where f_0 is the emitted frequency; similarly, ω_R equals $2\pi f_R$, where f_R is the received frequency. Since f_0 is fixed, equation (2) states that the frequency f_R of the sound reflected by the moving target and received at R differs from that of the emitted sound by an amount depending on the rate dr/dt at which the range is opened or closed. Stated differently, motion of the target causes the rate of variation of phase and thereby the period of the received sound to differ from that of the emitted sound. When dr/dt is constant, λ is independent of time and may be treated as a constant in differentiating equation (1). Also, when dr/dt is very small compared with c, the velocity of sound in the medium, λ may be set equal to either c/f_0 or c/f_R without significantly affecting the numerical results obtained from equation (2).

Substituting V for dr/dt, $2\pi f_0$ and $2\pi f_R$ for ω_0, and ω_R, c/f_0 for λ, and rearranging reduces the previous equation to

$$\delta f = \pm \frac{2fV}{c}, \tag{3}$$

where δf is the doppler shift $f_R - f_0$. The shift will be of positive sign—in other words, the received frequency will be higher than the emitted frequency—when the range is closing, and of negative sign when the range is opening. When δf is positive, the shift is termed "up-doppler"; when negative, "down-doppler."

The restriction that the motion of the target be perpendicular to OR, that is, in the direction of the acoustic axis of the echo-ranging transducer, is unnecessary. If the target T is actually moving with the velocity V, as shown in Figure 2, at an angle θ to the acoustic axis, then

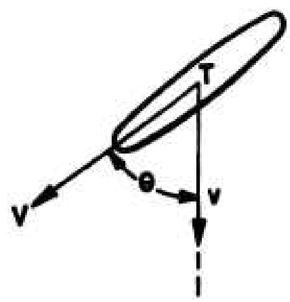


FIGURE 2. Components of target velocity.

the preceding argument applies to the component of motion along the acoustic axis, namely v, which equals $V \cos \theta$. Clearly, the component of motion parallel to the face of the transducer will not produce a doppler shift. Hence, equation (3) should be generalized to read

$$\delta f = \pm \left(\frac{2f_0}{c}\right) V \cos \theta. \tag{4}$$

Obviously, the preceding analysis applies equally well to the case of a stationary target and a moving transducer. Similarly, it applies when target and echo-ranging vessel are both in motion. Furthermore, it will be clear that the discussion leading to equation (4) requires only minor modification when the emitted sound is not a sustained tone but is instead a CW pulse of length τ. In this case, the essential spectrum of the pulse consists of a group of components extending over a band Δf, which is $2/\tau$ cycles wide, and each of these components will experience a frequency shift given by equation (4). If the pulse is very short and its essential spectrum correspondingly broad, the essential spectra of the emitted pulse and the echo reflected from a moving target will overlap each other unless the doppler shift δf is equal to the essential width Δf of the pulse spectrum. Thus, the minimum condition which must be met in order that the essential spectra of pulse and echo will not overlap is

$$\frac{2}{\tau} < \left(\frac{2f_0}{c}\right) V \cos \theta,$$

or

$$\tau > \frac{c}{f_0} \frac{1}{V \cos \theta}. \tag{5}$$

When τ is less than one-half the limit given in equation (5), resolution of the two spectra becomes very difficult.

Usually the frequency f_0 of the emitted pulse is in the supersonic region and for present purposes may be considered equal to 24 kilocycles. The received sounds are heterodyned to the region of audio frequencies. Since heterodyning merely subtracts a constant number of cycles from each component in the received sound, the magnitude of the doppler shift in cycles per second is the same in the heterodyned sounds presented to the ear as in the unheterodyned sounds in the water. Thus, use of a supersonic transmission frequency f_0 followed by heterodyning is advantageous because the doppler shift in the sounds presented to the ear is much greater than could be obtained by emitting a CW pulse at a low sonic frequency.

Since most of the scatterers producing reverberation may be considered as stationary, or nearly so, there are usually only two significant sources of doppler shift: motion of the echo-ranging vessel and motion of the target. The former shift, called *own-doppler*, is also given by equation (4), if V is understood to represent the speed of the listening vessel and θ the bearing of the transducer relative to the bow. Thus, the center frequency of the reverberation spectrum received when the transducer is trained forward will exceed that of the emitted pulse by $2f_0V/c$; and the reverberation frequency will be less than the transmitted frequency by the same amount when the transducer is trained aft. In these cases, θ equals zero or $-\pi$, and $\cos\theta$ equals ± 1. When the transducer is trained abeam, θ equals $\pi/2$, and $\cos\theta$ vanishes. In this case the center frequency of the reverberation spectrum coincides with that of the emitted pulse.

10.1.1 Magnitude of Doppler Shift

It will be useful to evaluate orders of magnitude for a number of doppler shifts. For example, the number of cycles of shift due to the radial velocity of a target opening or closing the range at a rate of 1 knot may be computed from equation (4). This calculation gives the pitch difference between target echo and reverberation received on a given bearing. The reverberation will in general have various amounts of own-doppler, but, unless the pitch of the reverberation is stabilized by some means such as *own-doppler nullifier* [ODN], the operator must judge the existence of target doppler by comparing the pitch of the echo and the reverberation. If we assume that the component of target motion $V\cos\theta$ in the direction of the transducer axis is 1 knot, or $1\frac{2}{3}$ feet per second, the doppler shift amounts to about 17 cycles for an echo-ranging frequency of 24 kilocycles. The sound velocity c has been set equal to 4,800 feet per second in this computation. Thus the doppler shift per knot of target range rate is 17 cycles at 24 kilocycles.

Since an echo-ranging vessel may often move at speeds of 20 knots or more, giving rise to positive and negative own-doppler shifts correspondingly larger than that computed here, the band width of an echo-ranging receiver operating at 24 kilocycles, and without ODN, should be at least 40 times 17 cycles, or 680 cycles. Otherwise, the received frequency of an echo with a relatively small amount of doppler may lie outside the cutoff frequencies of the receiver band, and may become so attenuated as to be inaudible.

It should also be mentioned here that own-doppler in the reverberation produced by pings emitted from a moving antisubmarine vessel is quite marked when heard aboard a submerged submarine. The submarine sonar operator first hears the directly transmitted pulse. The emitted power is high, the distances of interest fairly small, and even when the projector is trained away from the submarine, the sound intensity is at least within about 30 decibels of its value on the axis; hence the directly transmitted pulse can always be heard aboard the submarine. The received intensity of this pulse will, of course, vary with the orientation of the search projector during the pinging cycle. Immediately following the ping, the submarine operator hears the associated reverberation, produced largely by inhomogeneities toward which the projector of the surface vessel is oriented. If the projector is pointed toward the submarine, the reverberation will have the same pitch as the original pulse, except for frequency shifts resulting from the submarine's motion through the water. When the projector is pointed away from the submarine, the direct signal, radiated from one of the minor lobes of the projector, has the same frequency as before, but the reverberation will be sound scattered out of the main sound beam over to the submarine and may have quite a different frequency.

Suppose, for example, that the submarine is on the surface vessel's beam. The direct signal is therefore received at the frequency f_0 without doppler (except for that resulting from the submarine's own motion). If the projector is pointed forward, the sound received by the water and scattered as reverberation will have a frequency greater than f_0 by the amount Vf_0/c, where V is the speed of the surface vessel in feet per second. When this sound reaches the submarine it still has this higher frequency. Similarly, when the projector is pointed aft, the reverberation heard aboard the submarine will have a lower pitch than the directly transmitted pulse. By correlating variations of pulse intensity with the doppler shift of the ensuing reverberation, it is in principle possible to obtain tactically useful information on the speed and course of the echo-ranging vessel, provided that the surface vessel carries out a regular search plan, sweeping over a fixed sweep sector at a uniform rate.

10.1.2 Width of Reverberation Spectrum

The effect of pulse length on the reverberation spectrum has been discussed in Chapter 7. When the echo-ranging projector is in motion relative to the scatterers, the presence of own-doppler provides an added complication. Since own-doppler varies with the relative bearing and since the main lobe has an appreciable width, the reverberation received from different directions will have different doppler shifts. As a result, the spectrum of the received reverberation will have added broadening.

The magnitude of this effect may be computed without difficulty. Consider the rays in the main lobe which differ in direction from the acoustic axis by some angle $\Delta\theta$, measured in the horizontal plane. Let the resultant change of own-doppler from its axial value be denoted by $\Delta(\delta f)$. Since $\Delta\theta$ is a small quantity for standard echo-ranging gear, $\Delta(\delta f)$ may be computed by means of a Taylor expansion of equation (4) in which only the first term need be considered. This process yields

$$\Delta(\delta f) = \frac{d}{d\theta}(\delta f)\Delta\theta = \pm\left(\frac{2f_0}{c}\right)V \sin\theta\,\Delta\theta\,. \quad (6)$$

Hence, the spread in the reverberation spectrum produced by own-doppler will be greatest when the projector has a beam orientation, and θ equals $\pi/2$ or $3\pi/2$. When the projector is trained forward or aft, and θ equals 0 or $-\pi$, $\Delta(\delta f)$ found from equation (6) vanishes. Higher order terms in the Taylor expansion give a nonzero value for $\Delta(\delta f)$, but this value is generally so small that it may be neglected for most purposes.

The spread in the reverberation spectrum given by equation (6) for a projector trained abeam is quite appreciable. If the width of the major lobe of the hydrophone is 6 degrees on each side of the axis, giving a value of about 0.1 radian for $\Delta\theta$, and if the speed is 20 knots, the width of the reverberation spectrum resulting from finite beam width and own-doppler amounts to 70 cycles, since $\Delta(\delta f)$ found from equation (6) is about 35 cycles. This spread must be added to that resulting from the finite width of the emitted pulse spectrum. The center frequency of the reverberation spectrum received when the transducer is oriented toward 90 or 270 degrees coincides, of course, with that of the center frequency of the emitted pulse spectrum. However, it may be noted that the echo received from a stationary target on which the transducer axis is not quite centered will show a doppler shift relative to the reverberation. This shift also is produced by own-doppler and the finite width of the main lobe; it is entirely independent of the existence of relative target velocity and makes difficult the determination of target speed by means of doppler.

It is clear that the spread in the reverberation spectrum may depend on a variety of factors other than that associated with the length τ of the emitted pulse. Nevertheless, when τ is 10 milliseconds or less, the essential spectrum of the emitted pulse will be at least 200 cycles wide. Resolution of the spectra of reverberation and echo will be difficult, unless target doppler is very large. Equation (5) indicates that for 6 knots of target motion, which is fairly high, amounting to a doppler shift of about 100 cycles at an echo-ranging frequency of 24 kilocycles, the essential spectra of the echo and reverberation will overlap if the pulse length is less than 20 milliseconds. Resolution would probably be possible for somewhat shorter pulse lengths, but especially in view of the broadened reverberation spectrum produced by own-doppler, a pulse length of 10 milliseconds or less would make identification of target doppler difficult if not impossible.

10.2 EFFECT ON RECOGNITION

In a group of early British reports[1-4] describing tests under semifield conditions, evidence is presented on the effects which doppler may produce on echo recognition. From these data certain conclusions may also be drawn on the effects produced by changing the heterodyne frequency, the pulse length, and the injection range, and by substituting a loudspeaker for headphones. As indicated in earlier sections, precise definition of the various factors influencing test results is difficult even under laboratory conditions. The present tests were intended merely to serve as guides to performance which might be expected in the field with gear in use at the time the studies were made. While such an objective is extremely important operationally, it is limited in that it often fails to establish whether the controllable factors which determine the level of performance have been properly exploited.

10.2.1 Experimental Procedure

A standard echo-ranging transducer, mounted on the bottom, as is customary with harbor-defense installations, was used to produce the reverberations. Because of the shallow depth of the water, these were essentially bottom reverberations. The transducer was operated at a frequency of 15 kilocycles, and the received reverberation was heterodyned down to an audio frequency of 1 kilocycle, which is standard British practice. In addition, 0.5 and 1.5 kilocycles were also used as audio frequencies in some of the tests. The signal injected into the reverberation was produced by a 15-kilocycle oscillator provided with a calibrated vernier condenser by means of which the oscillator frequency could be varied from 14.9 to 15.1 kilocycles in 10-cycle steps. Thus, the frequency of the audio pulse could be varied in either direction from that of the reverberation, and by amounts up to 100 cycles. The level of the injected pulse could be varied in steps of 3 decibels; and the mixture of pulse and reverberation was amplified and heterodyned to the audio region.

Pulses of controlled duration were obtained from the oscillator by means of a magnetic relay which was actuated when a pair of moving terminals made contact with a pair of stationary terminals. In this manner, the duration as well as the injection range of the CW pulse could be predetermined. Care was exercised to eliminate false cues produced by key clicks and transients. Internal evidence indicates that the envelope of the generated pulse was rounded rather than rectangular, so that spurious cues of the type discussed in Section 9.1.2 played no significant role in the tests under discussion. With the aid of a CRO, the duration of the injected CW pulse was adjusted to equal that of the emitted pulse which produced the reverberations. Pulse lengths of 10 and 70 milliseconds were used.

The amplifier used had a rather sharply tuned input circuit. Hence, its response, shown in Figure 3, produced a relative loss of approximately 6 decibels for injected pulses whose frequency differed from 15 kilocycles by 100 cycles. This discrimination of the amplifier against dopplered signals must be borne in mind when examining the test results, since these are stated in terms of the input level of the primaudible pulse. No corrections for gain changes due to the tuning characteristic of the input circuit have been introduced in Figures 5, 6, 7, 8, and 10 since the illustrated characteristic refers to sustained tones and may require further modification before it can be applied to pulses.

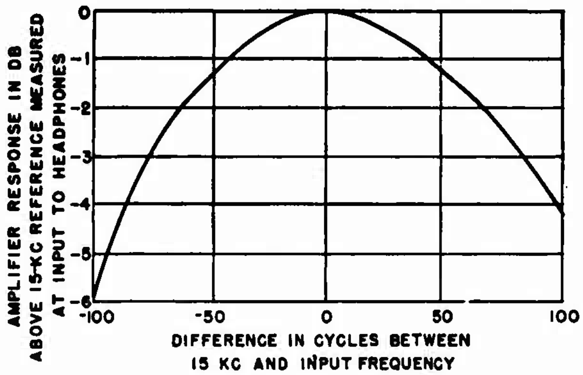


FIGURE 3. Frequency response of the test amplifier.

HEADPHONES

The response characteristics of the headphones and loudspeakers used are of importance in evaluating the results. Unfortunately, no information is available concerning the loudspeakers, but the threshold curve for the headphones, reproduced in Figure 4, is worth examining. The points shown were obtained with three observers wearing each of three different headsets in turn. These headsets were of the type used in the masking tests. The observed levels of just audible tones at each frequency are plotted in decibels relative to 1.3×10^{-6} volt. The rms amplitude of the faintest 1-kilocycle tone that could be heard was 4×10^{-6} volt.

Use of sustained tones was necessary in determining the headphone threshold, since key clicks and transients made results obtained with pulses unstable and unreliable. Since the echo masking tests were not performed with sustained tones, the results shown in Figure 4 may not be directly applicable. The level of room noise during the headphone study was somewhat higher than desirable and is believed to account for much of the scatter shown by the individual threshold determinations. However, the headphone tests extended over a period of several days, and the mean curve drawn through the points was considered reasonably free of the effects produced by extreme variations in the level of room noise.

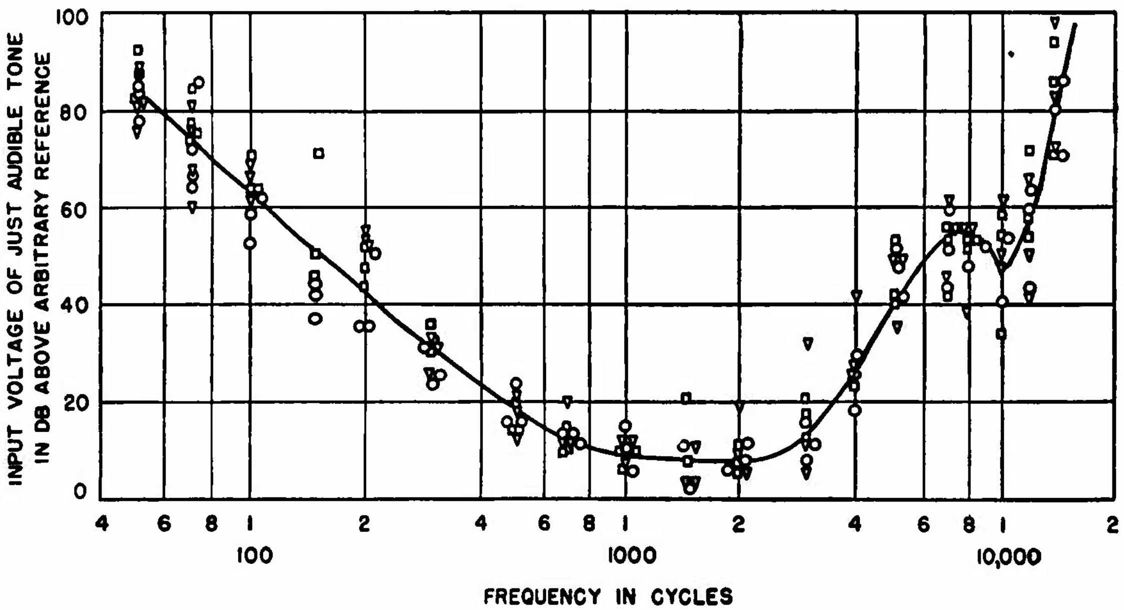


FIGURE 4. Power of sustained tonal signals just audible at various frequencies, for the headphones used in the reverberation tests.

The threshold shown in Figure 4 was obtained by matching the headphone impedance at 1 kilocycle to that of the signal generator. Since this procedure, which is the one followed in practice, neglects the possible effects of mismatch between headphone and generator impedances at other frequencies, a supplementary series of tests was made, in which nearly constant current was fed to the phones for a given generator setting at any frequency. This is in contrast to the nearly constant voltage supplied by the first arrangement, but the altered procedure had practically no effect on the shape of the resultant threshold curve. It is not known, however, whether the response of these headphones is linear with input; in other words, whether their relative response at various frequencies is the same at comfortable listening levels as at the threshold of audibility. The mean voltage input to the headphones for the reverberation used in the masking tests, stabilized by means of automatic volume control as explained later, was 7×10^{-2} volt. Since this corresponds to a sensation level of 85 decibels [20 log $(7 \times 10^{-2}/4 \times 10^{-6})$], which is somewhat higher than ordinarily considered comfortable, it may be inferred that the output of these headphones is not quite linear with input.

The threshold curve shown in Figure 4 agrees, in the region below 1 kilocycle, with that shown in Figures 76 through 79 in Chapter 4, although in both cases the threshold curve is much steeper than that computed on the basis of pressure in the ear canal. This departure is probably due to acoustic leakage at the low frequencies. In addition, however, the threshold shown in Figure 4 fails to drop in characteristic fashion between 1 and 4 kilocycles, indicating that headphone response was somewhat unfavorable in that region; and this indication should be borne in mind when assessing the relative degree of masking produced by reverberation on signals with various amounts of up-and down-doppler (see Figure 5, for example). In the region of 10 kilocycles, the threshold curve in Figure 4 shows evidence of a resonance. Since this resonance frequency varied slightly among observers and for a given observer by changing the manner in which the headphones were held to the ears, there seems little doubt that the effect was due to a cavity resonance in the region of the outer ear.

Volume Control

Immediately following a transmission, the level of reverberation background is high. This level usually declines in a fluctuating manner over an interval of 2 to 3 seconds, depending on oceanographic and operating conditions, until the received background consists of ambient or self-noise. To prevent overloading of the receiver and detector by the early blast of reverberation and yet have a gain setting high enough so that noise-limited echoes will be presented at a favorable level, use of a level stabilizer is required.[5-11] At least three forms of stabilizer are in common use. The system used by the British is a modified type of automatic volume control and, in the tests under discussion, held the output level of the reverberation nearly constant to a range of about 2,500 yards.

This is achieved by applying the average rectified output to the control grid of the first amplifier tube. In order to avoid application of bias while an echo is being received and the total level of the received sound is raised, application of bias is delayed for 100 milliseconds, which exceeds the duration of most echoes. Removal of bias is made to occur within 10 milliseconds after the received level begins to fall, so that there will be no discrimination against an echo which follows a reverberation blob or which follows another echo. Owing to the 100-millisecond delay in application of bias, protection against the initial high level of reverberation is secured by an independent biasing arrangement and automatic volume control takes over about 100 to 200 milliseconds after the transmission is completed.

The automatic gain control arrangement is designed to cease operating 6 seconds after a transmission and to resume only after another transmission. The need for this feature is imposed by the fact that such automatic control would reduce the degree of output modulation in propeller sounds which the sonar operator may wish to hear. Similarly, it would reduce the contrast between sounds received on the hydrophone axis and slightly off its axis.

Time-varied gain [TVG] and reverberation-controlled gain [RCG] are alternative forms of gain control. In the case of TVG, high bias is applied initially and allowed to taper off exponentially. Thus, no drop of gain can be produced by a high level blob just prior to an echo. With RCG systems, as with TVG, gain may increase but never decrease; however, the rate of gain restoration in the case of RCG is determined by the received level of the background. Thus, when the reverberation level falls below that of noise, full gain is restored.

Test Administration

The tests were administered individually to between 2 and 6 observers most of whom were inexperienced. It was found, however, that after a few trials they gave consistent results which were in close agreement with the performance of experienced listeners. Headphone presentation was used in all cases except those described in Section 10.2.6, in which the object of the tests was to compare performance obtained with headphone and loudspeaker presentation.

The level of primaudible pulses was determined by means of an ascending series type of minimal increment test. Observers had no foreknowledge of the range at which the signal was to be injected nor of the degree of doppler. These factors were held constant while the signal level was brought up to primaudibility from a point well below threshold. When the observer had satisfied himself that he could detect no signal at a given level, gain in the signal channel was increased by 3 decibels, and the test administrator noted the level of the signal finally identified as primaudible in three out of five trials. To discourage guessing, each observer was required to indicate the range and, if possible, the direction of the doppler shift imparted to the signal.

"Only one of the five observers who were asked to estimate the difference in frequency, if any, between a 70-millisecond [pulse] and the reverberations was able to do so with any accuracy. Three of the other four observers could distinguish a 20-cycle frequency change and state correctly whether it was high or low, but the remaining observer experienced great difficulty in deciding if the echo pitch was high or low, even when the frequency difference was as great as 60 cycles."[1] This variability in pitch discrimination (or rather, pitch identification) was apparently not strongly associated with the cue the observers used in these tests to identify the presence of an audible signal, since the difference in signal threshold levels for the best and poorest observers under given conditions rarely exceeded 6 decibels and was more often nearer 3 decibels. Furthermore, it is probable that some of the variability in recognition differentials was produced by the fact that the reverberation samples presented to successive observers differed significantly among themselves. More specific information on ability to identify doppler was obtained in a different set of tests.[3,4]

After signal threshold levels had been determined for each of the observers, the range, as well as the pulse frequency, was changed and thresholds were determined for the new condition. In this manner, the masking curves shown in Figures 5, 6, 7, 8, and 10 were determined. Each of the curves in these figures represents an independent test series. Slight differences in trends among these figures were produced by changes in the characteristics of the reverberation received from day to day during the tests. However, the data shown in any one figure were obtained on the same day, and over as short a time interval as possible, in order to minimize the effects due to changes in the received reverberation. It may be assumed therefore, that the effect of the chief variable indicated in each of the figures is relatively uninfluenced by instability of the masking background. This assumption is strengthened by the fairly high degree of consistency among the figures.

The lower horizontal scales in Figures 5 through 10 represent the difference between the frequencies of the pulse and the reverberation. Thus, the positive values refer to up-doppler or closing-doppler, and the negative values to down-doppler or opening-doppler. The upper horizontal scale in Figure 5 translates this frequency difference into units of relative target velocity, on the assumption that the emitted pulse has a frequency of 24 kilocycles [see equation (3)]. The vertical scale indicates the level of the primaudible pulse with the indicated degree of doppler. All the pulse threshold levels are expressed in decibels below an arbitrary reference, which was the same in all the tests. Thus, the figures do not give recognition differentials directly, but only changes in *RD* associated with doppler, pulse length, range, or heterodyne frequency. However, estimates of recognition differentials which may be compared with those shown in Figure 12 of Chapter 9, are given in Section 10.2.3.

10.2.2 Recognition Differentials

Figure 5 shows the effect of doppler in decreasing the level of the primaudible signal.

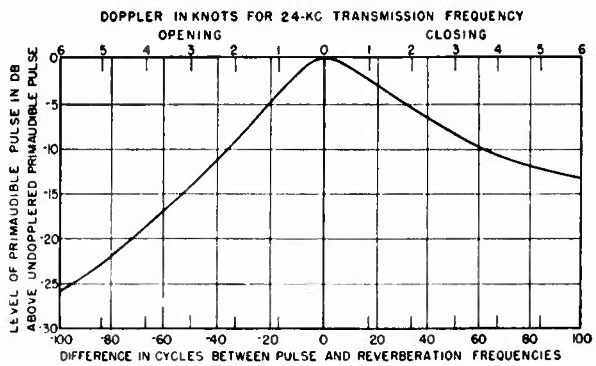


FIGURE 5. Effect of doppler on the detectability of pulses masked by reverberation.

The curve represents the mean performance of six observers and was determined in a relatively short time interval in order to offset effects produced by variation of the received reverberation with time. Range and bearing were fixed in this series of tests and the transmission as well as the signal had a duration of 70 milliseconds. The heterodyne frequency was set so that the presented reverberation had a center frequency of 1 kilocycle.

"The unsymmetrical character of the curve is most noticeable and the effect was readily demonstrated by introducing [a pulse] 100 cycles higher than the reverberation frequency and reducing it in intensity until it was no longer audible. Then without altering the intensity, the [pulse] was plainly heard when it was reduced in frequency to 100 cycles lower than the reverberation frequency."[1] In examining this curve, allowance should be made for the fact that it was obtained with nearly rectangular pulses, which correspond essentially to beam-aspect echoes. In practice, submarine echoes will not show significant target doppler unless they are obtained at bow, stern, or quarter aspects. Under these circumstances, the echo envelope is generally inferior to what is observed at beam aspect, and the target strength also deteriorates for aspects other than beam. Similarly, the reverberations used in the masking tests were produced and received by a stationary transducer; but, as pointed out previously, the width of the reverberation spectrum is a function of lobe width, ship speed, and hydrophone orientation. Consequently, the reverberation spectra received in many cases of practical interest will be significantly wider than those which produced the masking in the present set of tests.

However, even when allowance is made for the effects of these factors, the final conclusions will depend ultimately on the ear's performance in the case of dopplered and undopplered echoes, as indicated in Figure 5. Such conclusions have considerable significance, since they form the basis for rules of procedure guiding prosubmarine and antisubmarine operations; it is therefore desirable to examine the auditory results with some care.

Perhaps the broadest basis for such an examination is to compare the reverberation masking data with the results of tests for the masking of one sustained tone by another, as shown in Figure 2 of Chapter 2. In making such a comparison, it should be borne in mind that the sounds used in the present tests were not pure tones. The background had a spectrum of finite width, and its time-amplitude pattern showed rapid variations of amplitude and phase. Similarly, the signal had a duration of only 70 milliseconds. Thus, the essential widths of signal and reverberation spectra were about 29 cycles, and this is less than 50 cycles, or the width of a critical band centered at 1 kilocycle. However, such sounds would be expected to give substantially constant stimulation of the basilar membrane over a frequency interval corresponding to a full critical band and to produce progressively less stimulation at more distant frequencies, as found in other cases of remote masking. It seems significant, therefore, that the width of the peak in Figure 5—as measured to points where the threshold level of the dopplered pulse is 6 decibels more favorable than for the undopplered—is about 59 cycles, extending from about 23 cycles of down-doppler to about 36 cycles of up-doppler.

In general this same peak width, approximately that of a single critical band, would be expected for all cases in which the durations of emitted and received pulses exceed 70 milliseconds, which is nearly twice the reciprocal of the critical band width at 1 kilocycle. Conversely, the peak width should be greater than that of a critical band when the pulse is shortened sufficiently so that its essential spectrum extends over an interval wider than a critical band, or, in other words, when the pulse length is less than about 70 milliseconds and the audio frequency is about 1 kilocycle. This dividing line occurs at about the pulse length where the *RD* for undopplered pulses becomes relatively independent of their durations (see Figure 11 of Chapter 8 and Figure 14 of Chapter 9), and the inferred broadening of the peak, to a value of about 200 cycles for a 10-millisecond pulse, is indicated by the data in Figure 6.

It will be clear, therefore, that the minimum condition for the resolution of echo and reverberation spectra derived in Section 10.1 may not be applied unless these spectra exceed the width of the stimulated critical band, or, in general, exceed the resonance width of the affected receiver element. Thus, tests of the kind under discussion provide an additional method of determining critical band widths. Furthermore, the indications are that doppler shifts less than ± 18 cycles (which amounts to one-half the width of the narrowest critical band, and also to about 1 knot of relative target motion for an echo-ranging frequency of 24 kilocycles) cannot be detected by ear unless, perhaps, the audio frequency is reduced to about 200 cycles.

For such low presentation frequencies, subjective harmonics are readily generated, and separation of the segments of basilar membrane stimulated by the harmonics is greater than for the directly presented sounds. This process should be even more effective for lower heterodyne frequencies; but echoes with amounts of down-doppler exceeding the audio frequency would then emerge at the image frequency and might even appear to have up-doppler. Another obvious disadvantage of a low beat frequency is that abrupt shifts of phase associated with the irregular envelope of reverberation would become more noticeable to the ear as intervals of silence, so that the presented background would acquire a rattling quality and might produce a deterioration in performance. With the periodmeter, this increased number of abrupt phase shifts with decreasing heterodyne frequency would show up as a broadening of the reverberation spectrum.

Conversely, the harmonics may be produced by means of a nonlinear circuit. For example, a square law circuit, such as is used in the doppler doubler, will multiply the frequencies of signal and background by a factor of two, thereby doubling any frequency difference between them. However, such a mechanism, like the ear, will also produce the difference and the sum of the original frequencies. The sum frequency lies between the doubled values of the signal and background frequencies and, especially when the signal is relatively weak, will tend, since it obviously cannot be filtered out, to obliterate the doubling effect, giving a sensed doppler no greater than available without the doubler. This difficulty can be avoided by recording the incoming sounds and then presenting them to the operator by running the sound track at increased speed but introduces a new difficulty in that the active duration of the echo is diminished.

It should be explained at this point that the preceding estimate of ± 18 cycles as the practical lower limit for auditory determination of doppler shift should in general enable nearly certain identification of pitch differences, and that shifts of ± 9 cycles should be detectable 50 per cent of the time. Even this value, however, is considerably greater than the smallest pitch differences which can be detected between pure tones (see Figure 10 in Chapter 2). This deterioration in performance is due to the brevity of the pulse and to variations in pitch and intensity of the background. Thus, trained observers usually agree to within 3 cycles when they determine the salient frequency of a reverberation sample by matching against an oscillator tone, provided the final setting of the oscillator is taken as the average of those obtained by raising and also dropping the tone frequency to the apparent midfrequency of the reverberation. This procedure eliminates the effects of finite pulse duration and spectral impurity. Nonetheless, difficulty in obtaining a satisfactory match is experienced because the pitch of the reverberation fluctuates irregularly, and the observer must consciously try to find the midpoint of that zone. In addition, the mechanics of the basilar membrane sometimes make it difficult to distinguish between changes of frequency and of intensity, although training minimizes the effects of such illusions and biases. Because of the changing intensity of reverberation background with time, and also because many reverberations show systematic pitch drift, the most satisfactory reference for estimating doppler is that portion of the background immediately preceding the echo.

It may also be concluded that the pulse masking curve in Figure 5 shows only a single peak because the sounds used, unlike those on which the sustained tone masking data shown in Figure 2 of Chapter 2 are based, could establish no definite sensation of beats during the brief interval that the pulse was presented. However, as pointed out in Section 9.2.1, there are substantial fluctuations in amplitude due to phase interference when pulse and reverberation intensities are comparable. Since the intensity of a primaudible undopplered pulse, masked by CW reverberation, is about 7 decibels higher than that of the reverberation for a pulse length of 70 milliseconds, Figure 5 implies that the envelope of the pulse-reverberation mixture will not be substantially affected by interference except when signal doppler is ± 40 cycles and therefore the signal-to-reverberation ratio nearly unity. The fact that the greatest degree of masking occurs for the pulse whose intrinsic frequency is equal to that of the reverberation indicates that a 70-millisecond signal stimulates the same portion of the basilar membrane as does a sustained tone of equal frequency. In other words, effects associated with the time required for a stimulus to affect different points on the basilar membrane do not appear significant for signals of this length and intrinsic frequency.

While the threshold levels for pulses with small amounts of doppler are in good agreement with expectation, the fundamental significance of the relative thresholds for pulses with a doppler shift of 80 to 100 cycles is open to some question. There are several reasons for raising this question. One of these reasons is stated in the following paragraphs; the others are given in Sections 10.2.5 and 10.2.6, in connection with the data discussed in those sections.

The reason that the threshold levels for dopplered pulses do not decline abruptly is that remote masking occurs; that is, the finite spread of the disturbance on the basilar membrane produced by the reverberation renders dopplered pulses less audible than they would be in the absence of the reverberation background. In a general way, and aside from the absence of the dip that appears when the tonal signal has nearly the same frequency as the tonal background, Figure 5 resembles the masking curves obtained with sustained tones. This observation again implies that many of the results based on tests with sustained sounds remain at least approximately valid for the case of pulses (see also Section 8.1).

In fact, if the curve shown in Figure 5 is corrected for two known sources of bias, the computed recognition differentials for 0.9-kilocycle and 1.1-kilocycle pulses are very nearly in exact agreement with the recognition differentials which may be derived from Figure 2 in Chapter 2 for sustained 0.9-kilocycle and 1.1-kilocycle tones masked by a 1-kilocycle tone which has a sensation level of 80 decibels. This sensation level has been selected because the available data imply that the reverberation used in the British tests was presented at approximately this level. The two sources of bias referred to are (1) the characteristic of the tuning curve shown in Figure 3 and (2) the somewhat poorer response of the headphones for frequencies above 1 kilocycle, as indicated by the discussion of Figure 4. The computed recognition differentials for 0.9-kilocycle and 1.1-kilocycle tones masked by a 1-kilocycle tone are given in the middle column of Table 1, and the observed recognition differentials for a 70-millisecond pulse masked by 1-kilocycle reverberation are given in the right-hand column.

TABLE 1. Comparison between observed results and pure-tone data.

Tone or pulse frequency in cycles	*RD* in db for tone masked by 1-kc tone	*RD* in db for 70-millisecond pulse masked by 1-kc reverberation
900	−28	−23
1,100	−19	−14

In deriving the sustained-tone recognition differentials from Figure 2 of Chapter 2, the observations, originally expressed in terms of threshold shift, have been converted to decibels below the level of the masking tone. This has the effect of reducing somewhat the apparent degree of asymmetry between the masking of low- and high-frequency tones, because the level of the absolute audibility threshold falls in the interval between 0.9 and 1.1 kilocycles (see Figure 1 in Chapter 2). Also, it was necessary to interpolate among the observations made with an 0.8-kilocycle and 1.2-kilocycle masking tone in order to estimate the effect for a 1-kilocycle masking tone. These derived recognition differentials for sustained tones are uncertain by perhaps as much as 3 decibels. The curves from which they were read have been smoothed, and can probably not be read more accurately than this. In examining the tabulated pulse recognition differentials, it should be noted that Figure 5 gives the threshold level of the dopplered pulse in decibels below that of the primaudible, undopplered pulse, and, as indicated in

Figure 12 of Chapter 9, the latter was itself some 8 decibels above that of the reverberation.

In view of the uncertainties of measurement, and the assumptions required to arrive at the tabulated values, the agreement between the tone and pulse data is very satisfactory. Thus, the difference of 5 decibels between the two sets of recognition differentials is very nearly equal to the deterioration which would be expected to result from using a short pulse rather than a sustained tone.

It is also of interest to use the pure tone data for purposes of estimating the probable effect of listening level, which is given in Table 2. It may be concluded from Table 2 that recognition differentials for dopplered pulses will not be as favorable at low listening levels as at high. This conclusion is in agreement with the observation quoted in Section 9.2.2 for pulses without doppler, and is probably associated with the fact that smaller intensity changes can be detected at the higher listening levels. It will also be noted that the recognition differentials for 0.9-kilocycle and 1.1-kilocycle tones are more nearly equal at a listening level of 60 phons than at a level of 80 phons. The loss in

TABLE 2. Effect of loudness level on pure tone masking.

Tone frequency in cycles	*RD* in db for tone masked by 1-kc tone presented at 60 phons	*RD* in db for tone masked by 1-kc tone presented at 80 phons
900	—22	—28
1,100	—17	—19

audibility for the 0.9-kilocycle tone is somewhat greater than for the 1.1-kilocycle tone, presumably because the audibility threshold decreases with increasing frequency. Hence, the effective sensation level for the 0.9-kilocycle tone is lower than that for the 1.1-kilocycle tone. The loss of signal audibility of 2 to 6 decibels consequent on reduction of the listening level is of the expected order of magnitude (see Figure 17 in Chapter 9).

With American echo-ranging gear, the sounds are heterodyned to 800 cycles. For a masking tone of this frequency, the pure tone recognition differentials, computed from Figure 2 in Chapter 2, are listed in Table 3; so also are the results for masking tones with frequencies of 400 and 1,200 cycles. It will be seen from Table 3 that in general the recognition differentials are more favorable at the higher listening level and that the difference in recognition differentials between the low- and high-frequency signals is smaller at the lower listening level. It should be remembered that these values are uncertain by several decibels. If the pure tone data shown in Table 3 apply

TABLE 3. Recognition differentials for masking of tones by tones.

Frequency in cycles of masking tone	Frequency in cycles of masked tone	*RD* in db at 60 phons	*RD* in db at 80 phons
400	300	—33	—41
	500	—25	—25
800	700	—27	—31
	900	—18	—21
1,200	1,100	—16	—21
	1,300	—16	—17

even approximately to dopplered echoes about 100 milliseconds long and in the presence of reverberation, the indications are as follows. To begin with, dopplered echoes will be easier to detect at high listening levels and at low heterodyne frequencies. Secondly, the difference in the detectability of echoes with opening and closing doppler will be smaller at the lower listening levels. Consequently, level stabilizers offer the advantages of making it possible to listen at optimal values of the gain setting. Furthermore, the doppler data shown in Figure 5 are probably not reliable guides to performance for gear which is not provided with AVC or RCG, not only because the echo may arrive when the level of reverberation has fallen to a low value, but also because dopplered echoes may be threshold rather than masking limited if the gain is set too low. Finally, it will be clear why differences in the spectra of reverberation and rectangular undopplered pulses were able to depress the measured recognition differentials to negative values for the nonfiltered pulses described in Section 9.1.2.

10.2.3 Recognition Differentials without Doppler

During the tests described in the preceding section estimates were made of the primaudible signal-to-reverberation ratio for pulses without doppler by the following two methods. In the first, a CRO with recurrent time base and linear deflection scale was used to portray the instantaneous voltage developed across the headphones, and a visual estimate was made of the relative amplitudes of the primaudible signal and the reverberation at the range of pulse injection. In the second method, advantage was taken of the fact that the receiver was equipped with AVC, and the value of the control grid bias at the time of echo injection provided an indication of the relative levels of signal and reverberation at primaudibility.

Since the level of the reverberation background fluctuated over a wide range, in characteristic fashion, it was difficult to arrive at a very precise evaluation of the signal-to-reverberation ratio corresponding to primaudibility, but both methods indicated that the pulse became consistently audible when its amplitude was nearly double that of the neighboring reverberation peaks. This corresponds to an *RD* of 6 decibels for a 70-millisecond pulse without doppler (20 log 2). The report describing these tests[1] also remarks that a primaudible signal-to-reverberation ratio of 2, as indicated above, "is a conservative estimate; [pulses] of smaller amplitude may often be detected, possibly due to differences between the shapes of [pulses] and reverberation peaks." It should be noted at this point that the estimates of signal-to-reverberation ratio made by either of the methods used did not involve the difficulties usually attendant on visual detection of a signal masked by fluctuating background, since the test administrator knew the exact time of signal occurrence.

The *RD* of 6 decibels for an undopplered pulse 70 milliseconds in duration has been entered in Figure 12 of Chapter 9 as an open square; the same symbol is used to indicate the *RD* obtained in these tests for a 10-millisecond pulse without doppler, as described shortly. It will be seen that the squares are within 2 decibels of the line drawn through the filled-in circles which is well within the standard deviation of recognition differentials determined by the single-point method, and also well within the uncertainty of 3 decibels inherent in the present tests. Owing to the many systematic differences between these British tests and those at UCDWR, closer agreement would not be expected.

Comparison of signal threshold levels for durations of 10 and 70 milliseconds are shown in Figure 6. The data shown in this figure for

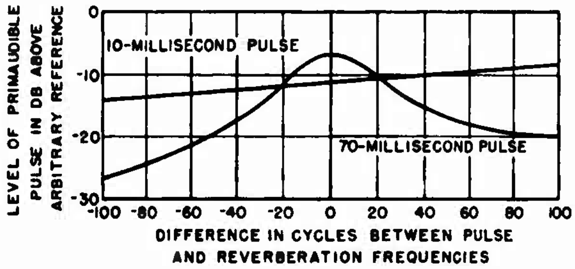


FIGURE 6. Effect of pulse length on the detectability of dopplered pulses masked by reverberation.

both pulse lengths were obtained within a relatively short time at a range of 500 yards and at fixed bearings. The bearing selected was not very critical in the case of the 70-millisecond pulse but, for the shorter transmission, the time-amplitude pattern and masking properties of the reverberation changed markedly with transducer bearing. In other words, echoes from small bottom features failed to overlap when the transmission length was made short enough, and the salient peaks in the received reverberation became difficult to distinguish from the signal, even when the latter had doppler. Consequently, the bearing used in the 10-millisecond tests was selected to give the smoothest reverberations obtainable.

If attention is confined to the threshold levels for 10-millisecond and 70-millisecond signals without doppler, it will be seen that the level of the primaudible signal was 4.5 decibels lower for the shorter pulse. Here, it should be mentioned again that all signal levels are stated in terms of the same arbitrary reference standard. Since the power output was independent of transmission length and since mean reverberation level for fixed conditions is inversely proportional to the length of the transmission, the level of the received reverberation was about 8.5 decibels (10 log 10/70) lower for the shorter transmission. However, detectability improved by 4.5 decibels when the 10-millisecond transmission was used; the assistance produced by the drop in reverberation level being offset in part by the reduced audibility of the shorter signal. Thus, the loss of audibility due to shortening the signal amounted to 4 decibels. In other words, the *RD* for a 10-millisecond signal is 4 decibels higher than that for a 70-millisecond signal, as shown by the squares and the filled-in circles in Figure 12 of Chapter 9. It should be noted that the listening levels were approximately equal for the tests at 10 and 70 milliseconds due to the AVC action of the receiver.

The slightly rising trend of signal threshold level with increasing pulse frequency which is shown by the 10-millisecond data in Figure 6 disappears when correction is made for the response characteristics of the input circuit and the headphones. Instead, the corrected curve shows an improvement of 6 decibels resulting from 100 cycles of down-doppler, and an improvement of 4 decibels for 100 cycles of up-doppler; the corresponding recognition differentials are 4 and 6 decibels, respectively. In other words, the width of the signal threshold curve is nearly 200 cycles between points 6 decibels down from the undoppleredsignal, and this is the theoretical width of the essential spectra of a 10-millisecond pulse and the reverberation produced by it. For amounts of doppler exceeding the essential widths of signal and reverberation spectra, resolution of the two should be possible; in other words, recognition differentials for doppleredechoes, as well as ability to identify doppler, may be assumed to improve with increasing doppler. This improvement, however, will be limited by the fact that the essential spectra are wider for the shorter pulses and also by the fact that appreciable amounts of energy exist outside the limits of the essential spectra. Both of these factors restrict the applicability of Table 3, so that it is most reliable for the longer pulses.

10.2.4 Effect of Range

Figure 7 shows that substantially the same results are obtained at pulse injection ranges of 500, 1,000, and 1,500 yards. The tests on

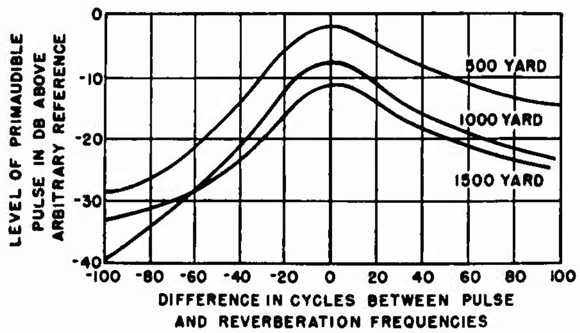


FIGURE 7. Effect of injection range on the detectability of pulses.

which this diagram is based were made over a relatively short time interval and with no more than 2 to 4 observers participating, in order to minimize effects due to change of reverberation level and character with time. The results are for a 70-millisecond pulse and a heterodyne frequency of 1 kilocycle.

The level of the primaudible undoppleredpulse is seen to diminish progressively with increasing range and in a manner corresponding to the progressive diminution of reverberation level with range. The average performance for all three ranges is in good agreement with the data shown in Figure 5; and the results for the three ranges agree among themselves to within 1 to 2 decibels except for opening-doppler in excess of 50 cycles, where the scatter is about 4 to 5 decibels. This scatter may perhaps result from the presence of noise masking. Signals with 100 cycles of down-doppler are likely to be masked at long range by noise components rather than by reverberation, since the primaudible echo in the presence of reverberation only is more than 20 decibels below the reverberation level and may therefore be weaker than the noise in a critical band.

It may be noted that some confusion in echo identification is to be expected at ranges where the level of the reverberation is about equal to the noise level in a critical band centered at the reverberation frequency, and where the reverberation is at the point of being masked by the noise. For example, if a sudden surge of reverberation level occurs under these circumstances, with the surge rising above the noise, the blob may easily be mistaken for an echo. No tests have been made under these conditions, but this point should perhaps be taken into account in training sonar operators.

10.2.5 Effect of Heterodyne Frequency

Since heterodyning subtracts a fixed number of cycles from the frequencies of the echo and reverberation components, the frequency difference between the two is not affected by this process. Hence, the ratio of the doppler shift to the reverberation frequency increases as the heterodyne oscillator is set to give the reverberation a lower audio frequency. To determine whether the value of this ratio has an effect on the recognition differentials of dopplered echoes, a series of tests was performed in which the heterodyne setting was selected to present the reverberation at frequencies of 0.5, 1, and 1.5 kilocycles. The test sequence covered a short period of time. Pulse duration was 70 milliseconds, and range and bearing were fixed. The observations are plotted in Figure 8.

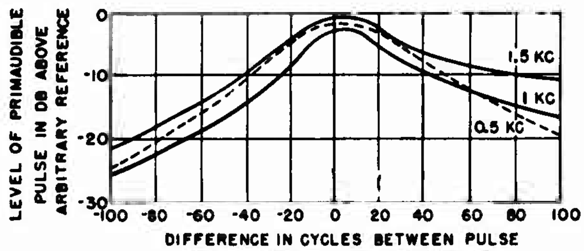


FIGURE 8. Effect of heterodyne frequency on the detectability of pulses.

The threshold level for an undopplered signal is seen to vary by no more than ± 1.5 decibels for the three frequencies tried, an amount less than the experimental uncertainty. This near identity is what would be expected from the fact that detection of undopplered signals is based largely on intensity discrimination, and this capacity does not depend strongly on frequency when the sensation level is fairly high (see Figure 15 in Chapter 2).

The results for the 1-kilocycle setting are in good agreement with the data shown in Figure 5. On the other hand, the effect of doppler, in the cases of the settings at 0.5 and 1.5 kilocycles, does not agree with expectation. Thus, the 0.5-kilocycle curve is more symmetrical and the 1.5-kilocycle curve less symmetrical than would be predicted from the sustained-tone recognition differentials shown in Table 3. This may mean that Table 3 is not a reliable indicator of pulse recognition differentials, since it does not take into account the spread of the reverberation spectrum outside the essential spectrum width Δf defined in Section 7.1. Although energy at these distant frequencies is negligible for most purposes, it may have some effect in the present case. The energy per cycle at any frequency f may be computed from equation (3) in Chapter 7. From this equation it is readily shown that the average energy per cycle in the spectral region between $\Delta f/2$ and $3\Delta f/4$ is about 23 decibels below the energy per cycle at the midfrequency f_0. Thus for a 70-millisecond pulse, with an essential spectrum width of 29 cycles, the average energy per cycle at a frequency of about 20 cycles (between 15 and 22 cycles) away from the midfrequency is 23 decibels down; in the neighborhood of 100 cycles the spectrum level is an additional 13 decibels down (20 log 100/20), or 36 decibels down in all. While this seems too low to explain the observed results, the uncertainty in the measurements is such that this effect cannot be ruled out.

In addition, the lack of improvement in the 0.5-kilocycle tests may be due, in part, to acoustic leakage around the headphone caps. Thus, although the same voltage was applied to the headphones in all cases, the listening level was probably very much lower in the 0.5-kilocycle tests than in the others; as shown in Table 2, a lower listening level may be expected to impair echo recognition. It is possible that the primaudible 0.5-kilocycle pulses were not far above the absolute audibility threshold. Furthermore, interference from room noise was probably more of a factor for the tests at 0.5 kilocycle, since most headphones do not efficiently insulate the ear from airborne sounds at the lower frequencies. Similarly, the enhanced asymmetry in the 1.5-kilocycle curve was probably caused, in part, by the unfavorable response of the headphones at frequencies above 1 kilocycle (see the discussion of Figure 4). Finally, it is not known to what extent system response was affected by changes in heterodyne frequency. It may be that, if proper corrections could be introduced for this factor, the shapes of the curves in Figure 8 would be in better agreement with Table 3, and also that the peak widths for the various heterodyne frequencies would conform more closely to the critical band widths.

It will be clear from the preceding discussion that recognition differentials for dopplered echoes masked by reverberation will be influenced much more by system response than is likely to be true for noise masking. This difference arises from the fact that, in the case of noise, the background components which produce masking are usually at frequencies in the immediate neighborhood of the echo frequency; hence, system response affects the levels of echo and masking components in much the same way. In the case of reverberation, on the other hand, dopplered echoes are subject to remote masking, that is, stimulation produced at points on the basilar membrane corresponding to frequencies higher and lower than that of the reverberation. This effect occurs within the ear; hence, a system which responds poorly at the echo frequency will be at a disadvantage because it discriminates against the echo but does not alter the amount of remote masking which takes place within the hearing mechanism.

The upper diagram in Figure 9 represents schematically A the auditory threshold; B a 1-kilocycle band of background noise admitted by a supersonic receiver and heterodyned to a center frequency of 1 kilocycle — this is shown as a shaded area and indicates the noise levels in 50-cycle bands; C the overall level of reverberation received in the same system at the same time and at a range of about 500 yards, which is represented as a vertical bar; and D the levels of primaudible echoes, 70 milliseconds long and with various amounts of doppler, shown as a dashed curve. The system is assumed to have a flat response, as shown by E at the top of the figure. For the sake of simplicity, the reverberation side bands are not represented. It is clear from this diagram that increased echo doppler tends to mitigate the effects of remote masking and thereby to improve recognition differentials. With increasing range, the reverberation level declines, and weaker echoes can be recognized. The latter condition is shown in the lower diagram, where everything is assumed to be the same as in the upper diagram, with the exception of the range or relative level of received reverberation. In this case, dopplered echoes are essentially noise masked; the effects of remote masking are negligible, and fainter echoes could be detected only if the level of received noise were reduced.

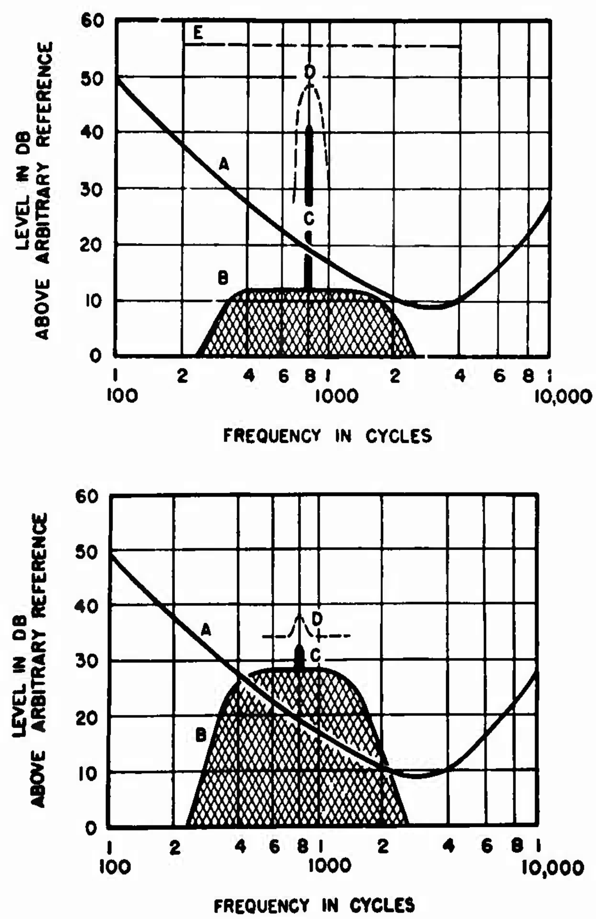


FIGURE 9. Effect of system response on the detectability of dopplered echoes.

The condition shown in the lower half of Figure 9 (level of reverberation presented to the ear not excessively high compared with background noise) could in principle be maintained at all ranges by using a notch filter, or a series of them, controlled by RCG. The notch filter is a bridged-T network, which can be made to discriminate very sharply against a narrow frequency band, without affecting response to frequencies removed only a few cycles from the limits of the notch. The depth of the notch, that is, the amount by which the level of the reverberation presented to the ear is attenuated with respect to the level of the received reverberation, can be controlled by an RCG circuit, in such a manner that the relations shown in the lower diagram are preserved at all times when appreciable amounts of reverberation are received. It would probably be desirable to leave the reverberation distinctly audible, to provide a pitch reference for doppler judgment.

Such a procedure would, of course, have no beneficial effect on the detectability of undopplered echoes. In fact, it would probably raise the recognition differentials for such echoes by introducing distortion and by dropping the stimulation level at the reverberation frequency. Its advantage for dopplered echoes, which may become increasingly important for the high-speed submarine recently developed, is obvious from the lower diagram in Figure 9.

In order to eliminate the effects of own-doppler, which would generally cause the frequency of the received reverberation to fall outside the limits of the notch filter, it would be necessary to rely on ODN or some other method of stabilizing the pitch of the received reverberation. ODN is an electronic device which samples the frequency of the received reverberation for a short interval immediately after each transmission. If the frequency of this short sample of reverberation is modified by own-doppler, the ODN unit alters the frequency of the heterodyne oscillator in such a manner that the output frequency of the reverberation at greater ranges is held to a constant value for a wide range of ship speeds and hydrophone orientations.

10.2.6 Effects of Headphone and Loudspeaker Presentation

These tests with headphones and loudspeaker presentation were undertaken in order to evaluate the significance of field reports received by the British which stated[4] that "weak echoes are sometimes detected on the loudspeaker when they are inaudible in the [headphones]." Similarly, it had been reported "that, with echoes of fair strength, doppler is readily noticeable on the loudspeaker, while at the same time no doppler is reported from the operator using the [headphones]."

Sixteen observers were tested on their ability to hear 70-millisecond pulses with various amounts of doppler in the presence of reverberation. The test procedure was the same as that described in the previous sections. Each observer was required to identify the level of the primaudible pulse, to state whether its frequency was the same as that of the reverberation, and, if not, to specify whether its frequency was higher or lower than that of background. In the present section, the discussion is concerned solely with the effect of loudspeaker and headphone presentation upon ability to detect dopplered and undopplered pulses.

Range and bearing were fixed. Each listener made fourteen observations, corresponding to seven doppler conditions (0, ±20, ±40, and ±60 cycles), repeated with loudspeaker and headphones. The amount of doppler was varied at random among the seven conditions stated, and the effect of variation in reverberation level with time was minimized by alternating frequently between loudspeaker and headphone presentation. Since the tests extended over a period of several days, the difference in performance obtained by the two presentation

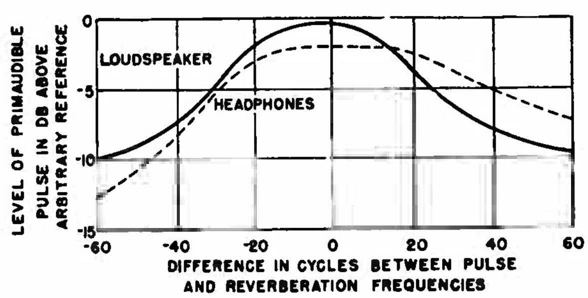


FIGURE 10. Effect of loudspeaker and headphone presentation on the detectability of pulses.

methods is much more significant than is the absolute level of performance. However, the latter, shown for both presentation methods in Figure 10, is in fair agreement with that shown in Figure 5.

The loudspeaker and headphones used were standard service types. In order to simulate field conditions, the room in which the tests were made was not quieted in any way. The airborne sounds were produced largely by people walking in and through the test room, and the level, though not the character, of this room noise was judged to be similar to that obtaining on the bridge of an antisubmarine vessel.

Figure 10 shows an advantage of about 1 to 3 decibels for headphone presentation in the range between 60 cycles of down-doppler and 10 cycles of up-doppler. For larger amounts of up-doppler, loudspeaker presentation improved performance by approximately 3 decibels. Thus, the test results confirm the indications obtained in the field.

The differences are not large, however, and are probably due to a combination of factors. It has already been indicated in Section 10.2.1 that the headphone response was somewhat unfavorable above 1 kilocycle. Furthermore, sensation levels and interference from room noise were probably significantly different for the two presentation methods.

10.3 UCDWR DOPPLER TESTS

Results of the UCDWR tests on reverberation masking of dopplered pulses became available too late to incorporate in and integrate with the preceding sections of this chapter. The observations and descriptive material upon which this summary is based constitute Part IV of a report[21] to be issued by the Sonar Data Division of UCDWR. Appreciation is expressed to the personnel of that organization for informally communicating these results in advance of publication. The same document may be consulted for further details relating to some of the informally communicated material described in Chapters 8 and 9.

10.3.1 Procedure

The test backgrounds were obtained from playbacks of film-recorded sea reverberations produced in deep water by 36- and 114-millisecond CW transmissions from standard, 24-kilocycle echo-ranging gear mounted on a surface vessel. The received reverberations were heterodyned to 800 cycles before recording, and all samples selected for use were free of significant pitch drift. The salient frequency of reverberation samples was established by means of measurements on the film recordings and checked by aural matching to an oscillator tone. Results secured by these two methods agreed, on the average, to within 1 cycle.

In all but one of the test series (see Section 10.3.4) the signal was a nearly rectangular pulse, either 42 or 118 milliseconds long, obtained from an oscillator tone by means of a gating circuit. These pulse durations were considered adequately matched to the reverberations produced by the 36- and 114-millisecond transmissions, respectively. Pulses with a fixed amount of "doppler" (determined by controlling the frequency of the oscillator in the signal channel) were injected at a fixed range. However, a number of different ranges were studied, all of them selected so that the signal occurred at a point beyond the initially overloaded section of the reverberation record, and several effects associated with this variable were established (see Section 10.3.3). Thus, some allowance can be made, when interpreting the present results, for the influence upon recognition of foreknowledge of the injection range. While no similar basis exists for discounting effects produced by foreknowledge of the amount and direction of doppler, the general similarity between the UCDWR and the British data, which was obtained by randomizing the doppler shift, indicates that such effects are probably fairly small.

The levels of successive pulses were varied at random among seven equally spaced values which extended over an interval of 14 decibels and which gave recognition probabilities varying from nearly zero to nearly unity. Use of blank presentations and other precautionary measures effectively eliminated commissive errors.

It may be mentioned in passing that difference in pitch between the signal and the reverberation context in which the signal occurs serves as an additional cue, so that the perceptual problem is somewhat more complex than in the no-doppler case. The momentary change of pitch during the life of the echo constitutes a type of auditory motion. Thus, the "pitch wobble" which characterizes most CW reverberation samples may occasionally result in multiple commissive errors when the observers' mental set is adjusted to doppler, while no such errors occur in no-doppler tests using the same reverberation background. Such commissive errors, however, belong to a different category from the usual ones, since they do not necessarily represent errors of judgment. Reverberation is composed of echoes from many discontinuities in the medium, some of them in motion. Probably as often as not there is no difference between the injected echo (or a target echo) and the constituent echoes of the background, other than the fact that one is wanted, and the remainder are not. The distinction can be made in the field by noting whether an echolike portion of the received sounds recurs at a given range, by listening for propeller beats, and so on.

Care was exercised in the design and operation of the test apparatus, as well as in the handling and storage of film, in order to minimize the influence of extraneous noise. Such effects were further reduced by transmitting the signal-background mixture through a band-pass filter with cutoffs at 565 and 1,130 cycles, which was sufficient to admit pulses with all degrees of doppler used in these tests. Harmonic distortion in the signal channel probably did not exceed 3 per cent for any condition studied.

The test sounds were presented to groups of five observers, the same group of five for any one test series, by means of high-quality headphones. All observers were young people, free of hearing defects; in preliminary tests with standardized material, they were found to rank with the highest 10 per cent of the population in pitch discrimination. For tests involving the effect of range (Figure 16), each doppler shift involved 300 judgments per observer, and for all other tests, between 1,000 and 1,500 judgments per observer.

At the instant of signal injection, the masking background generally had a sensation level of 55 to 65 decibels, and was at least 20 decibels above the threshold set by room noise. A small number of observations relating to the effect of listening level were made and are discussed below, but no absolute audibility thresholds were determined. Absence of information on audibility thresholds makes it impossible to express these data in terms of threshold shift and renders hazardous any attempts at precise comparison between the present results and either the British doppler observations or studies of the masking of one sustained tone by another.

Levels of signal and background were measured independently, at the output of the 565- to 1,130-cycle band-pass filter, in accordance with the methods described in Section 9.2. Except for Figure 18, all recognition differentials reported here were obtained by the single point method and probably are directly comparable to the no-doppler recognition differentials shown in Figure 12 of Chapter 9. No recognition differentials were determined for undopplered pulses in the present tests.

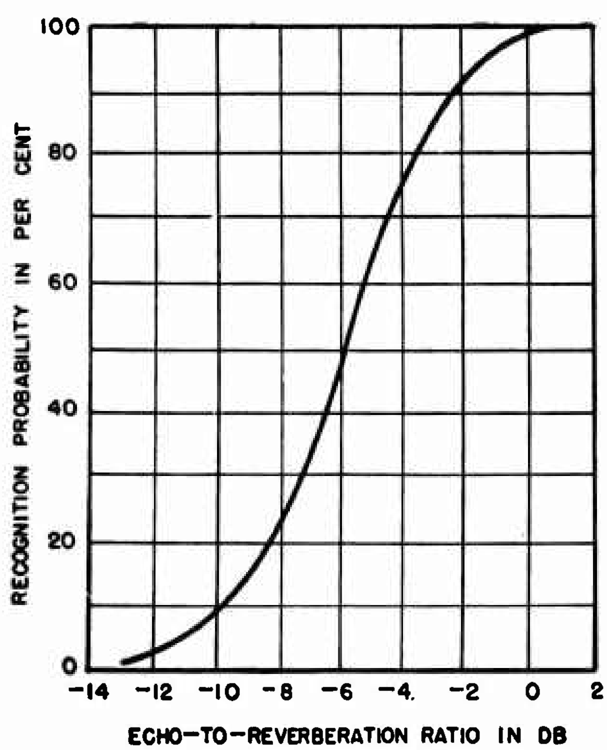


FIGURE 11. Composite transition curve for a series of five tests with a 118-millisecond pulse given 40 cycles of down-doppler.

Instead, it was assumed that the *RD* for the no-doppler case could be read directly from the data for recorded 36- and 114-millisecond beam aspect echoes given in Chapter 9, and the positions and shapes of the peaks of the curves shown in Figures 12, 13, and 16 were estimated in this way. Although the discussion in Section 9.1.2 indicates that this assumption may be questionable, the peaks of the curves in Figures 12, 13, and 16 form reasonable continuations of the experimentally determined portions. It is not certain, however, that the recognition differentials for dopplered pulses were unaffected by click transients (see Section 10.3.4).

Each of the recognition differentials shown in the various figures was determined from the 50 per cent point of a composite transition curve representing the performance of the group in a specific test series. A typical transition curve for a dopplered pulse masked by reverberation is shown in Figure 11. The shapes and slopes of transition curves for all degrees of doppler and both pulse lengths studied did not differ materially from those of the curve shown, and the latter is nearly an exact replica of the curve for a 114-millisecond beam aspect echo without doppler (see Figure 10 of Chapter 9). The standard deviations of recognition differentials for dopplered pulses were substantially the same as found for the no-doppler case discussed in Section 9.2.

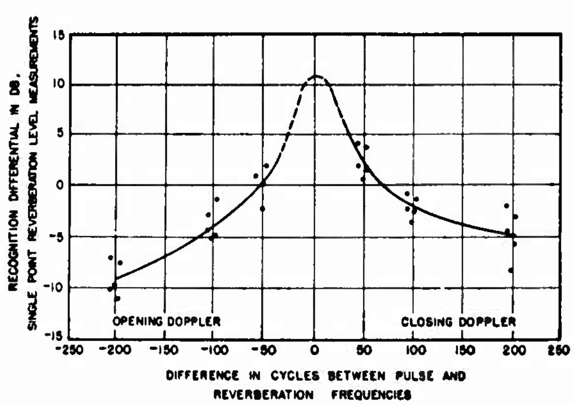


FIGURE 12. Effect of doppler on recognition differential for 36-millisecond pulses at a range of 1,400 yards. (Signal length equals 36 milliseconds; transmission length used in recording reverberation, 42 milliseconds.) Dashed portion of curve was estimated from data obtained in no-doppler tests with recorded echoes.

10.3.2 Effects of Signal Frequency and Duration

The dependence of recognition upon doppler and pulse length is shown in Figures 12 and 13. The masking curves in these figures extend over a frequency interval of 400 cycles and apply to pulse lengths of 42 and 118 milliseconds, respectively, injected at a range of 1,400 yards. In both cases, the reverberation film

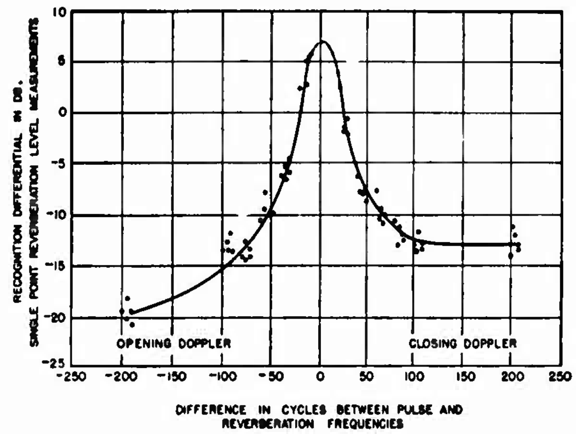


FIGURE 13. Effect of doppler on recognition differential for 118-millisecond pulses at a range of 1,400 yards. Dashed portion of curve was estimated from data obtained in no-doppler tests with recorded echoes.

loop consisted of five samples of background which had been carefully selected for uniformity of salient frequency. Because of small amounts of own-doppler and uncorrected changes in the setting of the heterodyne oscillator during recording, the average frequency characterizing a reverberation loop was not exactly 800 cycles; the average frequency of the five samples on one reverberation loop was 770 cycles, that of the other loop was 790 cycles. Similarly, the frequency difference between two reverberation samples on a loop never exceeded 10 cycles. This frequency scatter, as well as the scatter among recognition differentials associated with use of the single point method, is indicated by the circles in Figures 12 and 13. These points cluster in groups of five since they represent the results individually for each of the five reverberations on a loop.

The results of the British and the UCDWR doppler studies are collated in Figures 14 and 15. The following items should be borne in mind when examining these figures. First, the British data for 10- and 70-millisecond pulses apply to a reverberation frequency of 1,000 cycles presented at a listening level of about 80 decibels, whereas the UCDWR curves (nominally, for pulse lengths of 36 and 114 milliseconds, since those durations were used to produce the reverberation) represent a reverberation frequency of 800 cycles and a listening level of about 60 decibels. Headphone response was probably significantly different in the two sets of tests (compare Figure 4 with Figure 76 in Chapter 4). There is, in addition, some evidence that noise components in the recorded backgrounds affected the results of the UCDWR tests (see Section 10.3.3), but that such effects were absent or negligible in the British study (see Figure 7). Finally, it should be noted that, for the sake of removing the distorted appearance of the 10-millisecond results shown in Figure 6, corrections have been applied to compensate for the effects of restricted receiver band width and the audibility threshold (see Figures 3 and 4). The latter correction converts the 10-millisecond data into a masking audiogram, expressed in terms of threshold shift. Although it would be desirable to treat the remaining three curves in similar fashion, not enough of the required information is available.

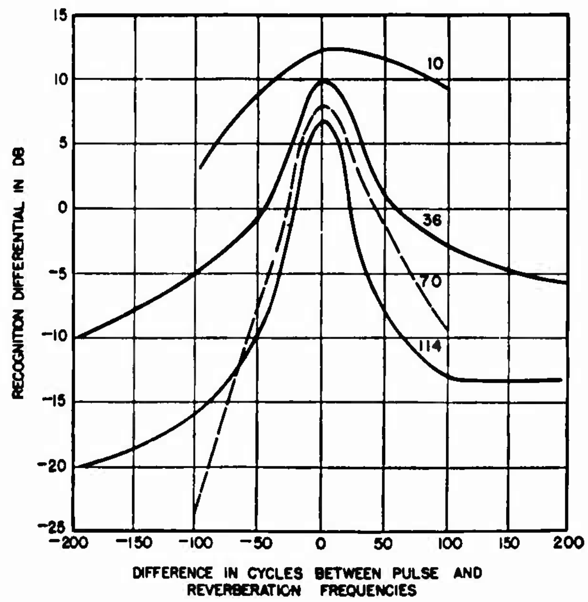


FIGURE 14. British and UCDWR masking curves showing effect of doppler on the detectability of pulses.

The recognition differentials for the no-doppler condition shown in Figure 14 have been read from the curve in Figure 12 of Chapter 9. These values determine the relative positions of the masking curves in Figure 14, which will be seen to form a reasonably consistent set. In agreement with the discussion in Sections 10.2.2 and 10.2.3, the peak widths of the masking curves tend to be equal to critical band width when the essential width of the spectrum is less than one critical band (70 and 114 milliseconds) and equal to the essential widths of pulse and reverberation spectra when these exceed the width of the critical band at the frequency in question. It will also be observed that large doppler shifts reduce the masking (change the *RD* relative to the no-doppler condition) by a greater amount for the longer pulse lengths. Both of these tendencies are to be expected from the fact that the widths of signal and reverberation spectra (including side bands) increase with diminishing pulse length; consequently, resolution of the two sounds becomes increasingly difficult, that is, it requires greater separation of their nominal frequencies, and the results of pure tone masking tests become progressively less relevant.

The shapes of the British and the UCDWR masking curves are significantly different, resulting in the intersection of the 70- and 114-millisecond curves. This difference in shape is probably due to the combined effects of differences in (1) listening level (see Table 3), (2) spectral characteristics of the noise and reverberation components in the masking backgrounds (see Section 10.3.3), and (3) pulse shapes used in producing the signal and reverberation.

To assist in evaluating the operational significance of the recognition differentials given in Figure 14, the curves have been replotted in Figure 15 to show the levels of the faintest audible echoes for a fixed set of conditions. Since the recognition differential equals

$$10 \log \frac{S}{B} = 10 \log S - 10 \log B \,,$$

where B and S refer to the intensities of the masking background and the primaudible signal, respectively; and, since B, in the case of reverberation background, is proportional to τ, the transmission length, it follows that

$$RD = 10 \log S - 10 \log k\tau,$$

where k is a constant characteristic of a particular set of operating conditions. Thus,

$$10 \log S = RD + 10 \log \tau + 10 \log k.$$

In other words, the ordinate in Figure 15, which represents the faintest detectable echo level, is derived by adding 10 log τ to the ordinate for each of the curves in Figure 14 (see also Figure 13 in Chapter 9).

It will be clear from Figure 15 that fainter echoes can be detected in the presence of reverberation for the longer pulse lengths, provided target doppler exceeds ± 25 cycles (about 1.5 knots for an echo-ranging frequency of 24 kilocycles); conversely, better results will be obtained in the case of stationary or slowly moving targets by reducing the pulse length. These conclusions refer specifically to auditory detection. They are further restricted at present by lack of detailed information relating to the effects of target aspect, of frequency and level stabilizers, of notch filters, and of heterodyne frequency.

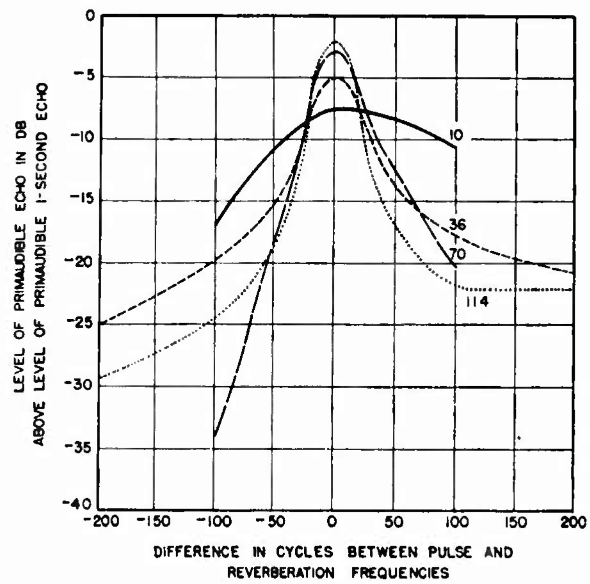


FIGURE 15. Estimated improvement resulting from changes in doppler and pulse length when reverberation is limiting.

10.3.3 Effect of Range

The effect of range on recognition of dopplered pulses was studied by using a single reverberation record produced by a 114-millisecond transmission. A 118-millisecond pulse with a fixed amount of doppler was injected at a fixed range, the only variable in a given test being the level of the injected pulse. Three test series were conducted, corresponding to ranges of 1,200, 1,600, and 2,000 yards. The recognition differentials obtained in these tests are shown in Figure 16.

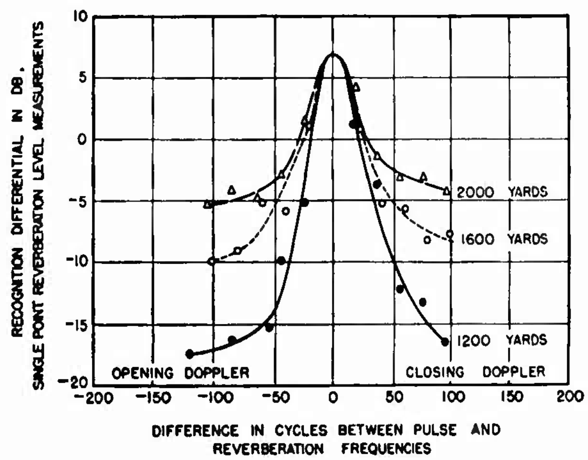


FIGURE 16. Effect of injection range on the detectability of dopplered pulses masked by reverberation.

Previous studies of the learning effects encountered in tests using a single reverberation sample as masking background indicated that such effects were probably substantially the same in all tests in this group; hence, the three curves in Figure 16 are directly comparable with each other. Preliminary observations indicate that when the injection range is unknown the recognition differential may be as much as 2 to 3 decibels higher than when the range is known in advance. This effect is greater for the large doppler shifts than it is for the small, being no greater, on the average, than about 1 decibel for undopplered signals (see Section 9.2.2). The larger effect associated with dopplered signals may arise from two circumstances. First, the perceptual situation is more complicated when the signal has a pitch different from that of the background (see Section 10.3.1). Secondly, the signal-to-reverberation ratio is relatively small when the pulse has a large degree of doppler; in other words, the background is much louder than the primaudible signal and more likely to divert attention from it.

For this particular reverberation sample, there is a marked dependence of recognition differential upon range. This dependence changes progressively with range, and is most noticeable for the larger amounts of doppler. It should be noted, however, that such range effects are observed in some cases but not in others (see Figure 7 and Section 10.3.4).

It seems reasonable to infer that the difference is due primarily to the presence of appreciable amounts of background noise at the ranges where these effects are observed. As indicated by the discussion of Figure 9, the character of the sound reaching the operator after a transmission will change progressively with time as the reverberation decays. During this decay, the relative intensities of the various background components are gradually altered, and the nature of the masking is likewise modified. Obviously, therefore, attempts to classify background types by listening may be misleading when one of the components of the mixture has especially striking properties.

The preceding discussion of the results shown in Figure 16 is supported by two other facts. First, for the larger doppler shifts the change of recognition differential with range is approximately equal to the change in level of the reverberation-noise mixture. In other words, as the background level falls, the remote masking of a dopplered pulse due to the reverberation becomes smaller than the adjacent masking produced on that pulse by the relatively more prominent noise components. Secondly, if this analysis is valid, the recognition differentials for pulses with large doppler injected at long range should be comparable with the recognition differentials for pulses masked by noise. The observed recognition differential for a pulse with 100 cycles of doppler injected at a range of 2,000 yards is, from Figure 16, about −4 decibels. Assuming that the measured background may be regarded as equally distributed within a band 565 cycles wide (which was the pass band of the filter used in making the measurements), but that the masking was produced exclusively by the noise components in a 50-cycle critical band centered at the pulse frequency, gives a critical band *RD* which is about 10 decibels (that is, 10 log 565/50) higher than that measured in the 565-cycle presentation band, or 6 decibels. This is in good agreement with the critical band *RD* of 7 decibels indicated by Figure 8 in Chapter 8 for a 114-millisecond pulse masked by noise.

It is also worth comparing the results for the 118-millisecond pulse given in Figure 13 with the results for the same pulse duration which are represented in Figure 16. The comparison shows that the curve obtained in the former case, for a range of 1,400 yards, falls midway between the curves for 1,200 and 1,600 yards in the second figure. In other words, pulses with large degrees of doppler in Figure 13 were apparently subject to some amount of noise masking, which may in part account for the intersection between this curve and the 70-millisecond curve in Figure 14. No plausible reason can at present be assigned for the constancy of the *RD* for up-doppler values between 100 and 200 cycles shown in Figure 13.

10.3.4 Detection of Recorded Echo

To add realism, a limited amount of data was obtained with a recorded 114-millisecond beam-aspect echo of constant intrinsic frequency. In this preliminary program the procedure which was followed (the only feasible plan in the time available) consisted of introducing doppler by using the single echo mentioned, as signal, and a reverberation background loop containing eight samples of different salient frequency and also produced by a 114-millisecond transmission. Of these eight, two had frequencies matched to that of the echo, two were of lower frequency, that is, the signal had different amounts of down-doppler with respect to them, and four gave various degrees of up-doppler. The playback from such a test loop resembles the sequence of reverberations heard aboard a moving antisubmarine vessel not equipped with ODN, when the axis of the hydrophone is trained to successive orientations during a search. It may be useful to simulate this practical situation more closely in further tests by varying the frequency of the signal as well as that of background. The total range of conditions tested in the manner described extended from 80 cycles of opening to 80 cycles of closing-doppler.

In order to reduce the effect of learning, since each sample on the background loop gave only one doppler condition, the echo was alternately injected at one of three predetermined ranges in each test. Unlike those already discussed, this group of tests showed no systematic dependence on range. Hence, the points shown in the results, graphed in Figures 17 and 18, represent averages for the three ranges used.

Since an *RD* determined by the single-point method for a single reverberation, even at three ranges, may have a low statistical reliability, the background levels were redetermined planimetrically by measuring over and prior to the echo. The recomputed recognition differentials,

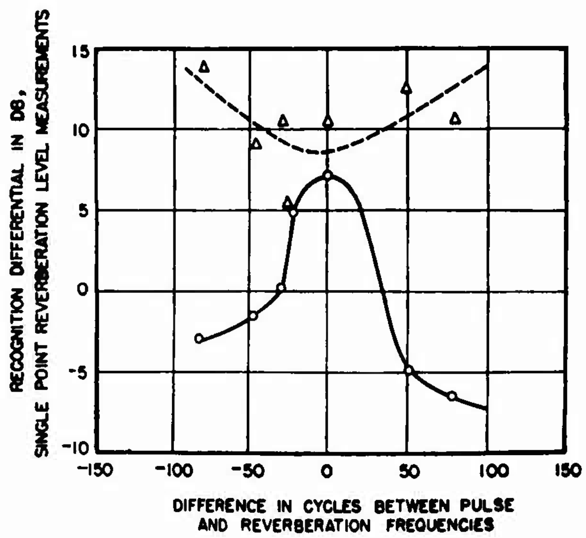


FIGURE 17. Effect of doppler upon detection of a recorded beam-aspect echo 114 milliseconds long. Doppler was produced by using recorded reverberation backgrounds whose pitch differed from that of the echo. Reverberation levels were measured by the single-point method; in the case of the dashed curve, the level of background was measured at the output of a 50-cycle band-pass filter centered at the echo frequency.

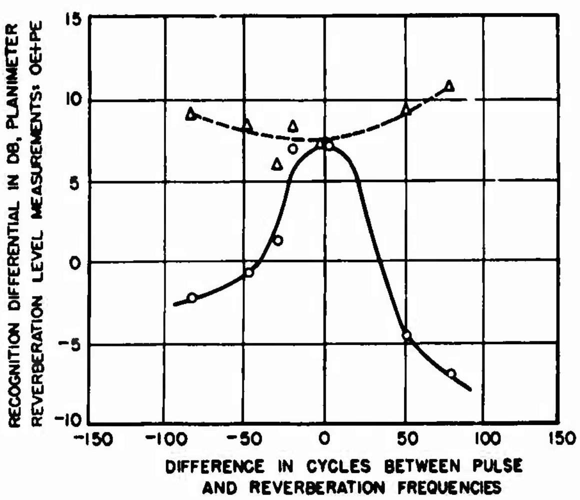


FIGURE 18. Effect of doppler upon detection of a recorded beam-aspect echo 114 milliseconds long. Doppler was produced by using reverberation backgrounds whose pitch differed from that of the echo. Reverberation levels were obtained by planimetry; in the case of the dashed curve, the level of background was measured at the output of a 50-cycle band-pass filter centered at the echo frequency.

shown by the circles in Figure 18, do not differ significantly from the single-point values, which appears to eliminate this factor as an explanation of the disagreement between the present and previous results.

The horizontal scale of these figures specifies the amount and direction of the frequency difference between echo and background, referred to the frequency of the background. The vertical scale gives the recognition differential obtained for each doppler condition.

The tests yielded three principal — and unexpected, or perhaps, anomalous — results. First, the curve defined by the circles in Figure 17, which gives the recognition differentials as determined by the single-point method, is asymmetrical in the opposite sense to those previously encountered. Secondly, doppler improves recognition much less than in the cases already described. Finally, in a few tests in which the listening level was raised by about 15 decibels above the 55- to 65-decibel level employed in the bulk of this series, the higher level impaired recognition by about 1 to 2 decibels.

Since each of the doppler shifts was associated with a different reverberation sample, while, in the other tests, each doppler shift was paired with, and the results averaged over, all of the samples, it was decided to investigate the possibility that the discrepancies arose from differences in the relative amounts of energy contained in the various samples at the echo frequency, that is, from differences in the shapes of their spectra. For this purpose, the reverberation level measurements were repeated using a 50-cycle, instead of a 565-cycle filter centered at the echo frequency, on the assumption that the masking was produced essentially by the background energy in a 50-

cycle critical band. The single-point and planimetric recognition differentials recomputed on this basi are given by the triangles in Figures 17 and 18 respectively.

It will be seen that planimetry reduces the scatter among the triangles, yielding a smoother curve. In addition, the circle and triangle for the no-doppler case coincide in Figure 18 but do not in Figure 17. The latter result is presumably due to the fact that the narrow filter rounds the reverberation envelopes, removing the more prominent amplitude peaks. In consequence, the single-point no-doppler recognition differentials based on measurements with a constrictive filter centered at the reverberation frequency would be expected to exceed those based on wide-band measurements. On the other hand, averaging the reverberation level over and prior to the echo should reduce the influence of envelope rounding and bring the wide- and narrow-band measurements into better agreement.

The dashed lines in Figures 17 and 18 curve upward from the no-doppler value. In other words, the ratio of primaudible signal to reverberation, both measured in a 50-cycle band centered at the signal frequency, is greater when signal and background differ in frequency. If this ratio were the same for all the frequency conditions tested it would follow that the process under consideration involves adjacent masking; that is, that masking arises primarily from the background components whose frequencies coincide with those in the essential spectrum of the signal. Thus, it may be concluded from Figures 17 and 18 that the masking is adjacent in the no-doppler case, but remote when significant doppler exists. Remote masking, however, is produced by the energy in the essential spectrum of the reverberation. Since this makes the dominant contribution to the level measured in the 565-cycle band, such measurements should provide an adequate basis for evaluating recognition differentials. Consequently, differences in the shapes of the reverberation spectra do not account for the departure of the present results from expectation. It may be inferred, incidentally, that the frequency discriminating mechanism of the ear does not have such sharp cutoffs as those of the 50-cycle filter used in these measurements.

If it could be demonstrated that the data given in Figures 17 and 18 are more representative of field conditions than those illustrated in Figure 14, that conclusion would have an immediate and obvious bearing upon choice of tactics. Thus, a submarine taking evasive action would be less detectable at a given speed and range when presenting its stern to a searching antisubmarine vessel than when presenting its bow. However, the preceding analysis does not by any means eliminate all the difficulties which must be removed before Figures 17 and 18 may be considered reliable.

It may be, for example, that the effect shown in these figures is real, that it will be met whenever the pitch of successive reverberations changes randomly, but that it will not be encountered in the absence of own-doppler shifts. To some extent, such effects are to be expected because the amount of remote masking produced by a background tone upon a signal which differs from it by a fixed number of cycles depends upon the frequency of the background (see Table 3), and this fact is ignored when, as in Figures 17 and 18, the frequency of background is treated as fixed.

Of course, the justification for following the procedure used in constructing Figures 17 and 18 lies in the fact that the range of variation of background frequencies is small. Hence, the masking curves should be much the same for all of them. Similarly, the effects of variation in auditory threshold, sensation level, and amplitude and frequency response of the test apparatus should be relatively small over the range studied. In addition, it seems improbable that any confusion produced by irregular changes in pitch among successive background samples could have any large effect on their ability to mask a signal, although it might be more difficult to make reliable judgments of the degree and direction of doppler under such circumstances.

Finally, it seems possible that the reason for the disagreement between Figures 14 and 18 is to be found in the properties (e.g., spectrum, envelope) of the recorded echo used as signal in the latter tests. In that case useful information may perhaps be obtained by repeating the variable reverberation pitch tests, and substituting a 114-millisecond pulse for the recorded echo. If this procedure proved helpful, further useful information might be obtained by modifying the pulse in various ways to simulate the commoner types of echoes. Clearly, it would be simpler to repeat the set of variable background items than it would be to select and control a new group of recorded echoes of variable pitch.

Chapter 11

SUMMARY OF ECHO STUDIES

11.1 BASIC FACTORS

As IN LISTENING to target sounds, a target echo can be heard only if it is sufficiently intense compared to the background of unwanted sounds. The echo level which can be heard half the time is called the *echo recognition level*. The difference in decibels between the echo recognition level and the background level, measured in some specified band, is called the *recognition differential*. In echo-ranging applications, a 1-kilocycle band is frequently used for specifying the background level.

11.1.1 Echo Characteristics

The following characteristics of echoes are significant in acoustic detection.

SPECTRUM

The sound power of the echo in each 1-cycle band is nearly proportional to the power of the outgoing pulse in the same band, except for a uniform doppler shift caused by the relative velocity of the echo-ranging projector and the target. For an ideal square-topped pulse τ seconds long, about 90 per cent of the energy is contained in a band $2/\tau$ cycles wide, centered at the nominal frequency of the pulse. Thus the "essential spectrum" of an echo from such a pulse is confined in a band $2/\tau$ cycles wide.

AMPLITUDE

Echoes from a small target or from a vessel at beam aspect will usually be very similar in shape to the outgoing pulse. However, echoes from targets which have appreciable extent in the direction of the sound beam may be considerably prolonged compared with the outgoing pulse, and their envelopes may be very irregular.

11.1.2 Background Characteristics

The background will generally include both noise — which may be airborne noise, electrical noise, or water noise — and reverberation produced by scattering of the pulse at the sea surface, sea bottom, and from scatterers in the ocean itself. The properties of noise are summarized in Chapter 6; for echo-ranging purposes, the spectrum level of noise in the water may usually be considered independent of frequency in the listening band. The properties of reverberation which are significant in echo recognition are summarized below.

SPECTRUM

For a stationary echo-ranging projector, the sound power per cycle in the received reverberation should be proportional to the power per cycle in the outgoing pulse. The phases of the different spectral components in the pulse and in the reverberation will be quite different, however. When the projector is moving, the change of own-doppler with relative bearing will broaden the reverberation spectrum unless the projector is pointed dead ahead or astern. Thus the essential width of the reverberation spectrum is $2/\tau$ cycles, plus an amount which depends both on the ship speed, the width of the main lobe of the transducer, and the relative bearing of the transducer. When a pulse of varying frequency is sent out, the reverberation spectrum covers the entire band of frequencies swept by the pulse, and doppler cannot be distinguished.

AMPLITUDE

The reverberation intensity (or amplitude squared) is directly proportional to the pulse length at any fixed range which is much greater than the pulse length. The reverberation intensity decreases with increasing range, and, with present gear mounted aboard a ship moving in deep water, usually falls below the noise level at a range between 1,000 and 2,000 yards.

The time-amplitude pattern of reverberation is very irregular. The average duration of a reverberation blob produced by a single-frequency pulse tends to be about the same as the durations of pulse and echo, thus adding to the difficulties of echo recognition. When a pulse of varying frequency is sent out, the duration of a reverberation blob becomes much shorter, while the length of the returning echo remains unchanged.

11.1.3 Echo-Background Mixture

The echo-background mixture should be presented at a loudness level in the neighborhood of 70 phons for best aural results. To achieve a constant loudness level, some form of variable amplification is required, since the reverberation intensity decreases with range. The listening band should be wide enough to pass echoes with any likely doppler shift, and, if very short pulses are used, should be at least $2/\tau$ cycles wide in order to pass the essential spectrum of the echo. Moderate limiting in the electrical circuit does not seem to have any significant effect on echo recognizability.

11.2 RECOGNITION LEVELS

11.2.1 Noise Background

Recognition differentials for 800-cycle pulses are given in Table 1. The recognition differentials are computed in terms of a 1-kilocycle band width. Narrowing the band width does not affect the pulse recognition level, provided the noise band width exceeds both the critical band width of the ear and the essential width of the pulse spectrum. Such a decrease in band width, however, increases the computed recognition differential, since the measured noise level decreases. The values in this table may be used for echoes from small objects and from ships at beam aspect and are probably valid for rounded as well as for square-topped echoes. For extended, irregular echoes, considerable deviations from these values may be expected.

TABLE 1. Recognition differentials for 800-cycle pulses.

Pulse length in seconds	*RD* in db for a 1-kc band
1 or more	—14
0.5	—12
0.2	— 9
0.1	— 7
0.05	— 4
0.02	0
0.01	4
0.001	16

For the shorter pulses, these values are also applicable to other heterodyne frequencies. At the longer pulse lengths, the recognition differentials are modified by the change of the ear's critical band width with frequency. Recognition differentials for long pulses of different frequencies, lasting 1 second or more, are given in Table 2. With mechanical or visual presentation, the recognition differentials shown in Tables 1 and 2 can be equaled only by use of a number of filters, similar to the critical bands of the ear. When recognition differentials for aural and visual presentation are about the same, a gain of about 2 decibels is obtained by use of both presentations simultaneously.

TABLE 2. Recognition differentials for long pulses of different frequencies.

Heterodyne frequency in cycles	*RD* in db for a 1-kc band
200	—14.6
400	—14.6
600	—14.6
800	—14.2
1,000	—13.8
1,500	—12.7
2,000	—11.5
4,000	— 8.0

REPEATED PULSES

Use of two pulses spaced about half a second apart gives an *RD* about 3 decibels lower (more favorable) than for a single pulse. If a string of pulses is sent out, successive pulses tend to merge into a steady tone if the pulse repetition frequency exceeds about 20 times per second. If the echo intensity is averaged over the duration of the pulse string and the noise level is measured in a 1-kilocycle band, the recognition differentials given in Table 2 for long pulses, or sustained tones, are applicable. For a repetition frequency less than 20 times per second, the recognition differentials computed with the peak echo intensity will be within 6 decibels of the values given in Table 1.

FREQUENCY-MODULATED PULSES

For a pulse at least 100 milliseconds long, the *RD* for a pulse of steadily increasing or decreasing frequency (chirp signal) is within 2 decibels of the *RD* for a constant-frequency pulse of the same duration. This conclusion presumably holds only if the frequency sweep does not extend much above 2 kilocycles, since at these higher frequencies increasing critical band width impairs performance.

11.2.2 Reverberation Background without Doppler

Recognition differentials for 800-cycle pulses, presented against reverberation of the same frequency, are given in Table 3. These recognition differentials are independent of the presentation band width, provided this exceeds the essential width of echo and reverberation spectrum.

TABLE 3. Recognition differentials for undoppleered pulses masked by reverberation.

Pulse length in seconds	*RD* in db relative to reverberation	Echo recognition level in db
0.5	2	0
0.2	5	—1
0.1	7	—2
0.05	9	—3
0.02	11	—5
0.01	12	—6

To estimate actual performance, the proportionality of reverberation intensity and pulse length must be taken into account. This improvement in performance is shown in the third column of Table 3, which gives echo recognition levels relative to the recognition level for a 0.5-second pulse.

The values of Table 3 are applicable to echoes from small targets and from vessels at beam aspect. Echoes which have highly irregular envelopes or which are prolonged relative to the outgoing pulse may show recognition differentials several decibels more or less than the values given in the table. Subject to this restriction, these values may be used for all heterodyne frequencies between 0.5 and 1.5 kilocycles; for frequencies outside this range, the validity of Table 3 is questionable. It may be noted that the simple types of visual and mechanical recognition would be expected to yield an *RD* less dependent on pulse length in the presence of reverberation, amounting to between 5 and 10 decibels.

For FM pulses presented against a reverberation background, the *RD* for aural recognition is probably in the neighborhood of 5 decibels less (more favorable) than for pulses of constant frequency and the same length, as long as the pulse length is not too short; at about 20 milliseconds, for example, the relative advantage of the frequency sweep probably disappears. Even lower echo recognition levels are probably obtainable by use of complex visual or mechanical systems with short FM pulses and suitable averaging and filtering circuits.

11.2.3 Reverberation Background with Doppler

With aural detection the recognition of echoes against reverberation is much facilitated by the doppler shift of echo relative to reverberation. For echoes and reverberation produced with a 70-millisecond pulse, the recognition differentials for different doppler shifts are given in Table 4.

TABLE 4. Recognition differentials for dopplered pulses masked by reverberation.

Doppler shift in cycles	*RD* in db above reverberation for a 70-millisecond pulse Up-doppler	Down-doppler
0	7	7
20	3	1
40	— 1	— 6
60	— 6	—13
80	—10	—18
100	—13	—24

Table 4 applies strictly only to 70-millisecond pulses injected in reverberation. The relative change of *RD* with doppler shift shown in this table is believed to be more generally applicable, however, since this depends only on the filter properties of the ear. Thus, for all echoes, even distorted ones of all pulse lengths greater than 70 milliseconds, the change of *RD* with changing doppler shift may be taken from Table 4. For shorter pulse lengths, however, the essential widths of pulse and reverberation spectra increase so much that the advantage of doppler is lost. At 10 milliseconds, for example, the *RD* is nearly independent of doppler for shifts up to 100 cycles.

The loudness level influences considerably the recognition differentials obtainable with dopplered echoes. Table 3 gives values for a uniform loudness level of 70 phons for the presented reverberation. For lower loudness levels, it is probable that the difference between up-doppler and down-doppler diminishes, and the recognition differentials for a fixed doppler shift becomes greater (less favorable).

Improvement of performance for dopplered echoes produced with a pulse length greater than about 50 milliseconds could presumably be obtained with a notch filter, with a cutoff sharper than that of the ear's critical bands. A notch filter would be required with visual or mechanical presentation to give performance comparable with that shown in Table 4.[a]

[a] N.B. Useful concepts, whose meaning is developed at diverse points in the preceding pages, have been indexed under unified headings to facilitate synopsis and review.

BIBLIOGRAPHY

Numbers such as Div. 6-510.12-M2 indicate that the document listed has been microfilmed and that its title appears in the microfilmed index printed in a separate volume. For access to the index volume and to the microfilm, consult the Army or Navy agency listed on the reverse of the half-title page.

Chapter 1

1. *Speech and Hearing,* Harvey Fletcher. D. Van Nostrand Company, New York. 1929.
2. *Hearing in Man and Animals,* R. T. Beatty. Bell, London. 1932.
3. *Theories of Sensation,* A. F. Rawdon-Smith. Cambridge University Press. 1938.
4. *Hearing,* S. S. Stevens and H. Davis. John Wiley and Sons, New York. 1938.
5. *An Introduction to Biophysics,* O. Stuhlmann, Jr. John Wiley and Sons, New York. 1943.
6. *Vibration and Sound,* P. M. Morse. McGraw-Hill Book Company, New York. 1936.
7. *A Textbook of Sound,* A. B. Wood. Macmillan Company, New York. 1941.
8. *Direct Observation of the Acoustic Oscillations of the Human Ear,* H. G. Kobrak. *Journal of the Acoustical Society of America,* Vol. 15, 1945, p. 54.
9. "Theory of the Mechanical Phenomena Occurring in the Internal Ear," J. A. Reboul. *Journal de Physique et le Radium,* Vol. 9, 1938, p. 185.
10. "Theory of Hearing, Form of the Vibration Pattern of the Basilar Membrane," G. v. Békésy. *Physikalische Zeitschrift,* Vol. 29, 1928, p. 793.
11. "Resonance in the External Auditory Meatus," N. Fleming. *Nature,* Vol. 143, 1939, p. 642.
12. "The Pressure Distribution in the Auditory Canal in a Progressive Sound Field," F. M. Wiener. *Journal of the Acoustical Society of America,* Vol. 18, 1946, p. 248.
13. "Resonance in the External Auditory Meatus," T. S. Littler. *Nature,* Vol. 143, 1939, p. 118.
14. "Altered Acoustical Resonance of the External Auditory Canal as a Factor in the Transient Improvement of Hearing by Fenestration Operations in Chronic Progressive Deafness," E. M. Josephson. *Journal of the Acoustical Society of America,* Vol. 13, 1941, p. 83.
15. "Aero-Otitis Media in Submarine Personnel," H. L. Haines. *Journal of the Acoustical Society of America,* Vol. 17, 1945, p. 136.

Chapter 2

1. "The Effect of High Altitude on Speech and Hearing," H. W. Rudmose, K, C. Clark, F. D. Carlson, J. C. Eisenstein, and R. A. Walker, *Journal of the Acoustical Society of America,* Vol. 18, 1946, p. 250.
2. "Experiments on the Pellet-Type of Artificial Drum," K. Lowy, *Journal of the Acoustical Society of America,* Vol. 13, 1942, p. 383.
3. "Immobilization of the Round Window Membrane, A Further Experimental Study," W. Hughson and S. J. Crowe. *Annals of Otology, Rhinology and Laryngology,* Vol. 41, 1932, p. 332.
4. "Reception of Sound by the Outer Ear," J. Tröger, *Physikalische Zeitschrift,* Vol. 31, 1930, p. 26.
5. "On Minimum Audible Sound Fields," L. J. Sivian and S. D. White, *The Journal of the Acoustical Society of America,* Vol. 4, 1933, p. 288.
6. "The Monaural Threshold, Effect of a Subliminal Contralateral Stimulus," J. W. Hughes, *Proceedings of the Royal Society (London),* Vol. 124B, 1938, p. 406.
7. "Fluctuation of Hearing Threshold," S. Lifshitz, *Journal of the Acoustical Society of America,* Vol. 11, 1939, p. 118.
8. "Piezoelectric Measurements on the Absolute Auditory Threshold for Bone Conduction," G. v. Békésy. *Akustische Zeitschrift,* Vol. 4, 1939, p. 113.
9. "Theory of Hearing, Sound Reception by Bone Conduction," G. v. Békésy, *Annalen der Physik,* Vol. 13, 1932, p. 111.
10. "Bone Conduction Threshold Measurements," N. A. Watson and R. S. Gales, *Journal of the Acoustical Society of America,* Vol. 14, 1943, p. 207.
11. "Cancellation of Cochlear Response with Air and Bone Conducted Sound," K. Lowy. *Journal of the Acoustical Society of America,* Vol. 13, 1942, p. 156.
12. "The Course of the Auditory Threshold in the Presence of a Tonal Background," E. G. Wever and S. R. Truman, *Journal of Experimental Psychology,* Vol. 11, 1928, p. 98.
13. "A Space-Time Pattern Theory of Hearing," Harvey Fletcher, *Journal of the Acoustical Society of America,* Vol. 1, 1930, p. 311.

14. "Mapping the Cochlea," E. A. Culler, J. Willman, and F. A. Mettler, *American Journal of Physiology*, Vol. 119, 1937, p. 292.

15. "Revised Frequency Map of the Guinea Pig Cochlea," E. A. Culler, J. D. Coakley, K. Lowy, and N. Gross, *American Journal of Psychology*, Vol. 56, 1943, p. 475.

16. "Differential Pitch Sensitivity of the Ear," E. G. Shower and R. Biddulph, *Journal of the Acoustical Society of America*, Vol. 3, 1931, p. 275.

17. *Hearing*, S. S. Stevens and H. Davis, John Wiley and Sons, New York, 1938.

18. "Differential Intensity Sensitivity of the Ear for Pure Tones," R. R. Riesz, *Physical Review*, Vol. 31, 1928, p. 867.

19. "Influence of Experimental Technique on the Measurement of Differential Intensity Sensitivity of the Ear," H. C. Montgomery. *Journal of the Acoustical Society of America*, Vol. 7, 1935, p. 39.

20. "Auditory Patterns," Harvey Fletcher. *Reviews of Modern Physics*, Vol. 12, 1940, p. 47.

21. "The Mechanism of Hearing," Harvey Fletcher. *Proceedings of the National Academy of Sciences*, Vol. 24, 1938, p. 265.

22. "Pitch, Quality, and Loudness of Musical Tones," Harvey Fletcher, *American Journal of Physics*, Vol. 14, 1946, p. 215.

Chapter 3

1. *Electrical Oscillations and Waves*, G. W. Pierce, McGraw-Hill Book Company, New York, 1920, p. 320.

2. *Explosions, Propagation in Linear Case* (Memorandum), Harry Nyquist, NDRC, Oct. 31, 1941. Div. 6-510.12-M2

3. *Binaural Listening System, Donald P. Loye*, NDRC 6.1-sr20-565, Interim Report P12/145, CUDWR-NLL, Jan. 12, 1943. Div. 6-621-M1

4. *Microstructure of Sea Water, Notes on Measuring Temperature Differences* (Memorandum), Harry Nyquist, C4-NDRC-045, Mar. 12, 1942. Div. 6-520.1-M5

5. *Subaqueous Listening, Directivity of a Pair of Rings* (Memorandum), Harry Nyquist, C4-NDRC-064, Apr. 2, 1942. Div. 6-560.2-M2

6. *Noise Trials of H. M. Submarines at Lochgoilhead, H. M. Sm. Tempest, N86, O. R. P. JASTRZAB (late American)*, OSRD W-233-1, Report AL/N.1/D.225, Admiralty Research Laboratory, Teddington, Middlesex, England, Apr. 2, 1942.

7. *Hydrophone Listening Tests* (Technical Memorandum), Donald P. Loye and Russell O. Hanson, NDRC C4-sr20-090, Report G12/2623, CUDWR-NLL, Apr. 20, 1942. Div. 6-625.2-M3

8. *Directivity with Two Microphones* (Memorandum), Harry Nyquist, C4-NDRC-071, June 12, 1942. Div. 6-560.2-M4

9. *Sea Background Noise*, OSRD WA-362.5, Report ARL/R4/D.225, Admiralty Research Laboratory, Teddington, Middlesex, England, July 1942.

10. *Directivity with Two Microphones, Addition vs. Multiplication of Outputs* (Memorandum), Harry Nyquist, C4-NDRC-073, July 1, 1942. Div. 6-560.2-M5

11. *Binaural Phenomena*, Ralph C. Maninger, Internal Memorandum P12/4089, CUDWR-NLL, Sept. 29, 1942. Div. 6-560.1-M2

12. *Binaural Listening System, Tests on Amada, September 15, 1942*, J. W. Horton, Donald P. Loye, Internal Memorandum P12/4021, CUDWR-NLL, Sept. 18, 1942. Div. 6-560.22-M1

13. *Binaural Tests on the AMADA, September 24, 1942*, Donald A. Proudfoot, Internal Memorandum P12/4090, CUDWR-NLL, Oct. 1, 1942. Div. 6-560.22-M2

14. *Effects of Airplane Noise on Listening with Headphones, Experiments Conducted at the Psycho-Acoustic Laboratory, Harvard University, September 27 to 29, 1942 and November 6 to 8, 1942*, William B. Snow and William D. Neff, OEMsr-20, NDRC 6.1-sr20-550, Navy Projects NS-106 and NS-198, Memorandum D16/D34/107, CUDWR-NLL, Jan. 25, 1943. Div. 6-580.31-M1

15. *The Use of the CK Tube as an Anti-Submarine Listening Device*, Irving Langmuir and E. F. Hennelly, General Electric Company, Feb. 15, 1943. Div. 6-622.1-M1

16. *Survey of Underwater Sound, Introduction*, Vern O. Knudsen, R. S. Alford, and J. W. Emling, 6.1-NDRC-729, Report 1, Feb. 26, 1943. Div. 6-580-M1

17. *Measurements and Analysis of Sound Pressures of Torpedoes in Range 40 cps to 128 kc/s*, A. B. Wood, OSRD WA-497-35, M/S Summary 6127/42, Mine Design Department, Leigh Park House, Havant Hunts, London, Eng., Mar. 10, 1943. Div. 6-913-M1

18. *Reduction of Range of Detectability of Submarines by Reason of Increase in the Depth of Submergence*, OSRD WA-897-25A, Report ARL/95.14/R.1, Admiralty Research Laboratory, Teddington, Middlesex, Eng., August 1943.

19. *Binaural Non-Rotating Directional Radio Sonic-Buoy System*, Elliott J. Lawton, OSRD 1878, NDRC 6.1-sr323-1108, General Electric Company Aug. 16, 1943. Div. 6-624.2-M2

20. *Submarine Noise and Performance of JP-1 Sonic Listening Equipment Aboard the USS Corvina, SS226,* [from] *September 13* [to] *14, 1943,* H. C. Williams, Navy Projects NS-113, NS-337, Memorandum for File D24/526, CUDWR-NLL, Sept. 17, 1943. Div. 6-580.1-M1

21. *Transmission of Noise Through JP-1 Training Gear,* Hollie C. Williams, Navy Projects NS-113 and NS-337, Memorandum for File D24/550, CUDWR-NLL, Oct. 14, 1943. Div. 6-623.1-M2

22. *Submarine Listening Systems, Report of Conference September 29, 1943,* C. R. Sawyer, Navy Projects NS-113 and NS-337, Memorandum for File D24/P30/560, CUDWR-NLL, Oct. 26, 1943. Div. 6-623.1-M3

23. *Survey of Underwater Sound, Sounds from Submarines,* Vern O. Knudsen, R. S. Alford, and J. W. Emling, NDRC 6.1-NDRC-1306, Report 2, Dec. 31, 1943. Div. 6-580.1-M2

24. *Key West Trials, January 10* [to] *16, 1944,* W. L. Clearwaters and P. E. Fish, Navy Projects NS-106 and NS-198, Memorandum for File D16/778, CUDWR-NLL, Mar. 2, 1944. Div. 6-580.1-M3

25. *Expendable Radio-Sono Buoy Listening Ranges, Second Key West Trials with USS Pintado, March 13* [to] *19, 1944, American Fleet-Type Submarine,* P. E. Fish, Navy Projects NS-106 and NS-198, Memorandum for File D16/843, CUDWR-NLL, Mar. 29, 1944.

26. *Auxiliary and Underway Tests on USS Gabilan, SS-252,* Donald P. Loye and Malcolm T. Rodger, Navy Project NS-212, Memorandum for File D52/866, CUDWR-NLL, Apr. 10, 1944. Div. 6-641.31-M1

27. *Oscillograms of 24-kc Noise Produced by a Destroyer,* George E. Duvall, Navy Project NS-140, Internal Report A2, UCDWR, May 1, 1944. Div. 6-580.2-M2

28. *Underwater Sound Measurements on Aircraft Carriers,* Navy Project NO-163. Memorandum for File M-212, Listening Section, UCDWR, May 15, 1944. Div. 6-580.2-M3

29. *Sound Output Measurements on the USS Bluegill at 62, 150, and 250-Foot Keel Depths,* NDRC 6.1-sr1046-1050, Navy Project NS-164, MIT-USL, May 15, 1944. Div. 6-580.1-M6

30. *Progress Report on Listening Systems for Patrol Craft,* OSRD 1043, NDRC C4-sr692-531, BTL, Nov. 12, 1942. Div. 6-622.1-M5

31. *A Study of Binaural Perception of the Direction of a Sound Source,* Irving Langmuir, V. J. Schaefer, C. V. Ferguson, and E. F. Hennelly, OSRD 4079, NDRC 6.1-sr323-1840, General Electric Company, June 30, 1944. Div. 6-560.1-M5

32. *Survey of Underwater Sound, Ambient Noise,* Vern O. Knudsen, R. S. Alford, and J. W. Emling, 6.1-NDRC-1848, Report 3, Sept. 26, 1944. Div. 6-580.33-M2

33. *Directivity at Low Sonic Frequencies,* Walter F. Graham and Ralph C. Maninger, Memorandum for File P33/1067, CUDWR-NLL, Nov. 8, 1944. Div. 6-612.21-M21

34. *Prediction of Sonic and Supersonic Listening Ranges,* OSRD 4761, NDRC 6.1-sr1131-1884, Navy Project NS-140, Sonar Analysis Section, CUDWR-SSG, December 1944. Div. 6-570.1-M6

35. *Sound Cavitation Tests on USS Springer, SS414,* Navy Project NS-140, Memorandum for File M-283, Listening Section, UCDWR, Dec. 20, 1944. Div. 6-580.1-M9

36. *Ambient Noise Survey, Miami Area and the East Coast of the United States,* Henry B. Hoff, Donald L. Cole, and Robert A. Wagner, NDRC 6.1sr1128-1944, Navy Project NO-163, Completion Report D46A/1215, CUDWR-NLL, May 5, 1945. Div. 6-580.33-M3

Chapter 4

1. *An Investigation of the Audibility of Underwater Noises,* W. F. Higgins, OSRD WA-317-18, Report 1, Physics Division, National Physical Laboratory, Teddington, Eng., July 17, 1942. Div. 6-560.21-M1

2. *An Investigation of the Audibility of Underwater Noises,* W. F. Higgins, OSRD WA-317-19, Report 2, Physics Division, National Physical Laboratory, Teddington, Eng., July 24, 1942. Div. 6-560.21-M2

3. *An Investigation of the Audibility of Underwater Noises,* C. G. Darwin, OSRD 2125-6B, Report 3, Physics Division, National Physical Laboratory, Teddington, Eng., Apr. 28, 1944. Div. 6-560.21-M3

4. *Submarine Listening Range at Loch Goil,* Rodney F. Simons, OSRD WA-1730-8D, Technical Memorandum, OSRD, London, Eng., Mar. 8, 1944. Div. 6-580.1-M4

5. *United States Fleet Anti-Submarine and Escort of Convoy Instructions,* Fleet Tactical Preparations Bulletin 223, Office of Commander in Chief, U. S. Fleet, Navy Dept., 1944.

6. *Masking Experiments* ([Part] I), NDRC 6.1-sr30-1751, Navy Projects NS-163 and NS-164, Report U-229, Listening Section, UCDWR, June 28, 1944. Div. 6-560.21-M4

7. *Masking Experiments* ([Part] II), NDRC 6.1-sr30-1757, Navy Projects NS-163 and NS-164, Report U-258, Listening Section, UCDWR, Sept. 15, 1944. Div. 6-560.21-M6

8. *Underwater Sound Measurements on Aircraft Carriers*, Service Project NO-163, Report M-212, Listening Section, UCDWR, May 15, 1944. Div. 6-580.2-M3

9. *Underwater Sound Output* [of] *USS Spot, SS413*, NObs-2074, Problem 2G, Report M-296, Listening Section, UCDWR, Mar. 1, 1945. Div. 6-580.1-M10

10. *Underwater Sound Output of USS Tinosa, SS283*, Problem 2G, Report M-303, Listening Section, UCDWR, Mar. 5, 1945. Div. 6-580.1-M11

11. *Survey of Underwater Sound, Sounds from Surface Ships*, M. T. Dow, J. W. Emling, and Vern O. Knudsen, OSRD 5424, NDRC 6.1-NDRC-2124, Report 4, June 15, 1945. Div. 6-580.2-M7

12. "Influence of Experimental Technique on the Measurement of Differential Intensity Sensitivity of the Ear," H. C. Montgomery, *Journal of the Acoustical Society of America*, Vol. 7, 1935, p. 39.

13. *Listening Systems for Patrol Craft (Final Report)*, NDRC 6.1-sr692-1698, BTL, Dec. 1, 1944.

14. *Interval Tests*, H. L. Bumbaugh, Memorandum for File D13/249, CUDWR-NLL, Apr. 8, 1943. Div. 6-560.1-M3

15. *Peak Recognition Tests*, Donald L. Cole and Sylvester J. Haefner, Memorandum for File P33/1053, CUDWR-NLL, Aug. 14, 1944. Div. 6-560.21-M5

16. *Airplane Detection by Submarines*, Edward Gerjuoy, Memorandum for File P33/897, CUDWR-NLL, May 10, 1944. Div. 6-580 31-M2

17. *Visual Detector of Torpedoes*, R. F. West and R. S. Rae, OSRD WA-506-17, Internal Report 110, HMA/SEE, Fairlie Laboratory, Eng., Dec. 12, 1942.

18. *Report of Conference* [Held at New London on] *March 10, 1943, on Listening Techniques*, William B. Snow, Report G1/242, CUDWR-NLL, Mar. 29, 1943. Div. 6-621-M2

19. *Phase Actuated Locator (Final Report)*, Reginald L. Jones, OSRD 1897, NDRC 6.1-sr695-997, BTL, Aug. 30, 1943. Div. 6-622.1-M2

20. *Response Characteristics of Interphone Equipment* (Revision IV to be inserted in OSRD 687), Francis M. Wiener, H. Wayne Rudmose, and others, OSRD 3105, OEMsr-658, Cruft Laboratory, Harvard University, Jan. 1, 1944. Div. 17-436.321-M3

21. *Headphone Comparison Tests*, Merritt B. Jones and William B. Snow, Memorandum for File G27/952, CUDWR-NLL, June 1, 1944. Div. 6-624.12-M5

22. *Analyses of Reduction Gear Noise in Water* [on the] *USS Blackfin, SS322*, Edwin F. Teal and Robert W. Pratt, Navy Project NS-212, Memorandum for File D53/1042, CUDWR-NLL, July 24, 1944. Div. 6-580.1-M7

23. *Permoflux Headset for JP Series Equipment*, William B. Snow, Navy Project NS-113, Memorandum for File D24/1128, CUDWR-NLL, Sept. 14, 1944. Div. 6-623.1-M6

24. *Frequency Weighting in Ship Sound Measurements, An Audibility Meter*, R. S. Gales, L. S. Goldberg, and A. M. Small, Report M-399, UCDWR, Mar. 19, 1946. Div. 6-560.21-M7

25. *Masking of Submarine Noise by a Sound Composed of a Finite Number of Single Frequency Tones*, Edward Gerjuoy, Report 09.421 (SAG-32), CUDWR-SAG, Apr. 28, 1945. Div. 6-560.2-M6

26. *Underwater Telephony*, J. W. Horton, OSRD 5183, NDRC 6.1-sr1128-2211, Navy Project NS-248, Completion Report D56/1415, CUDWR-NLL, May 15, 1945. Div. 6-623.42-M5

27. *An Experimental Study of Masking by a Line Spectrum*, Problem 2M2, Report M-314, NObs-2074, Sonar Data Division, UCDWR, June 7, 1945. Div. 6-560.2-M7

28. *Sonic Listening Aboard Submarines*, OSRD 5311, NDRC 6.1-sr1131-1885, Navy Project NS-140, Sonar Analysis Section, CUDWR-SSG, February 1945. Div. 6-623.1-M8

Chapter 5

1. *Listening Tests on May 14, 1943, of Hydrophone for Directional Sonic Buoy*, Ralph C. Maninger, Service Projects NS-330 and AC 55, Memorandum for File D34/349, CUDWR-NLL, May 19, 1943. Div. 6-624.21-M1

2. *Comparative Tests on Submarine and Surface Craft Listening Equipments*, Donald P. Loye and Ralph C. Maninger, NDRC 6.1-sr20-1020, Navy Project NS-113, Report D24/D38/391, CUDWR-NLL, Sept. 10, 1943. Div. 6-623-M4

3. *Bearing Accuracy of 3-ft. and 1-ft. Straight Magnetostriction Hydrophones*, Ralph C. Maninger, Memorandum for File D17/543, CUDWR-NLL, Oct. 9, 1943. Div. 6-612.62-M17

4. *Comparative Tests on 3-ft., 4-ft., and 5-ft. Hydrophones*, Ralph C. Maninger, Memorandum for File P33/949, CUDWR-NLL, June 12, 1944. Div. 6-612.62-M27

5. *Tests of JP-1 Equipment with 3-ft., 4-ft., and 5-ft. Hydrophones and QB Sound Gear*, W. F. Graham, Memorandum for File P33/1031, CUDWR-NLL, July 13, 1944. Div. 6-623.1-M5

6. *Comparative Field Tests of Underwater Listening Equipment Installed on the Fleabel*, Walter F. Graham and Ralph C. Maninger, OSRD 4391, NDRC 6.1-sr1128-1569, Report P33/862, CUDWR-NLL, Sept. 30, 1944. Div. 6-622.1-M3

7. *Tests of an Improved JP-1 Type Hydrophone on the USS Blueback*, Arthur L. Thuras, Navy Project NS-113, Memorandum for File P33/1161, CUDWR-NLL, Oct. 3, 1944. Div. 6-612.62-M38

8. *Experimental Investigation of Factors Involved in Sonic Listening*, Ralph C. Maninger, NDRC 6.1-sr1128-1932, Report P33/1319, CUDWR-NLL, Feb. 28, 1945. Div. 6-621-M7

9. *Equipment Developed and Used on the AMADA for Underwater Sound Investigations*, Walter F. Graham, Memorandum for File P33/1379, CUDWR-NLL, Feb. 28, 1945. Div. 6-621-M6

10. *Listening Tests of September 14, 1943, on Directional vs. Semi-Nondirectional Hydrophone*, Ralph C. Maninger, Memorandum for File G1/494, CUDWR-NLL, Sept. 28, 1943.

11. *Fundamental Listening Studies at the New London Laboratory*, Ralph C. Maninger, NDRC 6.1-sr1128-2210, Summary Report P33/1409, CUDWR-NLL, May 30, 1945. Div. 6-621-M8

12. *Prediction of Sonic and Supersonic Listening Ranges*, OSRD 4761, NDRC 6.1-sr1131-1884, Navy Project NS-140, CUDWR-SAG, December 1944. Div. 6-570.1-M6

Chapter 7

1. *Frequency Characteristics of Echoes and Reverberation*, W. M. Rayton and Raymond C. Fisher, OSRD 4159, NDRC 6.1-sr30-1740, Navy Project NS-140, Report U-244, UCDWR, Aug. 9, 1944. Div. 6-520.3-M2

2. *Submarine Detection, Directivity Indications* (Memorandum), Harry Nyquist, NDRC, Oct. 11, 1941. Div. 6-560.2-M1

3. *Reverberations in Sea Water, Theoretical Considerations* (Memorandum), Harry Nyquist, NDRC C4-NDRC-036, Dec. 19, 1941. Div. 6-520.1-M3

4. *Submarine Detection, Tentative Discussion of Reverberations* (Memorandum), Harry Nyquist, NDRC, Oct. 20, 1941. Div. 6-520.1-M2

5. *Submarine Detection, Air Bubbles in Sea Water* (Memorandum), Harry Nyquist, NDRC, Oct. 28, 1941. Div. 6-540.2-M1

6. *Submarine Detection, Aural vs. Automatic Reception* (Memorandum), Harry Nyquist, NDRC, Dec. 1, 1941. Div. 6-560.4-M1

7. *Submarine Detection, Scheme Employing Frequency Modulation in Transmitted Signal* (Memorandum), Harry Nyquist, C4-NDRC-039, Jan. 5, 1942. Div. 6-560.3-M1

8. *Binaural Listening for BDI Asdic Sets*, Lt. Com. F. Moller, R.Nor.N.V.R., T. M. Frey, and H. J. Hawkins, OSRD WA-4372-1K, Internal Report 214, HMA/SEE, Fairlie Laboratory, Eng., Mar. 27, 1945.

9. *Build-up Time of Selective Systems*, OSRD W-262-15, Internal Report 67, HMA/SEE, Fairlie Laboratory, Eng., Mar. 19, 1942.

10. *Detection of Radio Echoes, Noise Discrimination* (Memorandum), Harry Nyquist, C4-NDRC-063, Mar. 27, 1942. Div. 6-560.31-M1

11. *Reverberations in Sea Water, Note on Reflections at the Surface* (Memorandum) Harry Nyquist, C4-NDRC-065, Apr. 3, 1942. Div. 6-520.11-M1

12. *Split Beam Direction Finding, Interim Report*, OSRD WA-107-11, Internal Report 77, HMA/SEE, Fairlie Laboratory, Eng., Apr. 22, 1942. Div. 6-631.41-M2

13. *Reaction Times to Sight and Sound Under Certain Conditions*, OSRD WA-171-9, Internal Report 90, HMA/SEE, Fairlie Laboratory, Eng., June 1942.

14. *Observations of Echo Signals Obtained Using Variable Frequency Transmission* (Memorandum for file), Edwin M. MacMillen, NDRC C4-sr30-208, Navy Project NS-140, UCDWR, July 4, 1942. Div. 6-510.3-M1

15. *Echo Ranging at Lower Frequencies*, L. M. Langer, OSRD 1954, NDRC 6.1-sr30-1114, Navy Project NS-140, Interim Report U-109, UCDWR, Sept. 1, 1943. Div. 6-510.1-M2

16. *Hydrophone Effect from Icebergs*, G. E. R. Deacon, OSRD WA-1101-10, Internal Report 144, HMA/SEE, Fairlie Laboratory, Eng., Sept. 13, 1943.

17. *The Reverberation Equalizer*, George W. Downs, Jr., OSRD 3100, NDRC 6.1-sr30-692, Navy Project NS-142, Report U-97, UCDWR, Sept. 18, 1943. Div. 6-520.2-M2

18. *Echo Detection of Small Targets*, H. F. Willis, D. W. Boston, W. A. Jones, and R. T. Ackroyd, OSRD WA-1154-1b, Internal Report 145, HMA/SEE, Fairlie Laboratory, Eng., Sept. 23, 1943. Div. 6-633.24-M1

19. *Submarine Detection, Pinging Patterns* (Memorandum), Harry Nyquist, 6.1-NDRC-126, Mar. 28, 1944. Div. 6-560.3-M3

20. *Echo Ranging from Submarines, Range Indicator Requirements*, J. Warren Horton, Navy Project NS-142, Memorandum for File P29/1076, CUDWR-NLL, Aug. 16, 1944. Div. 6-631-M4

21. *Echoes from Swells*, George E. Duvall, Navy Project NS-141, Internal Report A43, UCDWR, Oct. 27, 1944. Div. 6-540-M1

22. *Echo Doppler Indicator*, J. Warren Horton, OSRD 4511, NDRC 6.1-sr1128-1924, Navy Project NS-142, Completion Report P36/1262, CUDWR-NLL, Dec. 15, 1944. Div. 6-631.34-M5

23. *Recommendations for Revision of Submarine Echo-Ranging System,* J. Warren Horton, OSRD 5273, NDRC 6.1-sr1128-2219, Navy Project NS-142, Summary Report P29/1435, CUDWR-NLL, May 24, 1945. Div. 6-631.5-M6

24. *Mine and Torpedo Detection Equipment,* J. Warren Horton, OSRD 5552, NDRC 6.1-sr1128-2220, Navy Project NS-297, Progress Report P66/1436, CUDWR-NLL, May 28, 1945. Div. 6-633.21-M2

25. *A New Method of Classifying Bottomed U-Boats by Means of Shadows,* P. C. Browne, E. Gresty, and H. Nodtredt, OSRD WA-4615-9, WA-5057-6, and WA-5176-11, Internal Report 220, HMA/SEE, Fairlie Laboratory, Eng., June 5, 1945.

26. *Sonar Hut Noise Measurements on a Frigate,* NObs-2074, Problem 2F, Report M-324, Listening Section, UCDWR, June 11, 1945. Div. 6-580.31-M3

Chapter 8

1. *Telegraph Signaling, Aural Reception vs. Other, Report of Laboratory Tests,* R. S. Alford, OEMsr-346, Report 3410-RSA-MS, Case 23 228-15, BTL, July 2, 1942. Div. 6-560.31-M2

2. "Mathematical Analysis of Random Noise," S. O. Rice, *The Bell System Technical Journal,* Vol. 23, 1944, p. 282, and Vol. 24, 1945, p. 46.

3. *Status Report on Task No. 5, Effect of Short Pulse Lengths and Receiver Bandwidth on Echo Ranging,* BTL, July 15, 1944. Div. 6-632.03-M5

4. *Frequency Modulation, Saw Tooth Modulation Equivalent to Succession of Single Frequency Transients with Rounded Envelopes* (Memorandum January 5, 1942), Kenneth W. Pfleger, NDRC, File 36680-3, BTL, Revised Mar. 19, 1942. Div. 6-635.1-M1

5. *Selectivity in Asdic Reception,* OSRD WA-362-9, Internal Report 109, HMA/SEE, Fairlie Laboratory, Eng., Nov. 9, 1942. Div. 6-560.3-M2

6. *Masking Effect of Water Noise on Short Pulses,* Raymond C. Fisher, NDRC 6.1-sr30-1538, Navy Project NS-221, Report S-239, UCDWR, July 25, 1944. Div. 6-634.2-M1

7. *The Detectability of Repeated Pulses,* Raymond C. Fisher and Carl F. Eckart, Navy Project NS-221, Memorandum for File M-261, UCDWR, Sept. 27, 1944. Div. 6-560.3-M5

8. *Submarine Detection, Directivity Indications* (Memorandum), Harry Nyquist, NDRC, Oct. 11, 1941. Div. 6-560.2-M1

9. *Reverberations in Sea Water, Theoretical Considerations* (Memorandum), Harry Nyquist, C4-NDRC-036, NDRC, Oct. 19, 1941. Div. 6-520.1-M3

10. *Submarine Detection, Aural vs. Automatic Reception* (Memorandum), Harry Nyquist, NDRC, Dec. 1, 1941. Div. 6-560.4-M1

11. *Telegraph Theory, Aural Reception vs. Other* (Memorandum), Harry Nyquist, NDRC, Dec. 11, 1941. Div. 6-560.4-M2

12. *Sensitivity of Recorder Paper,* OSRD WA-254-14, Internal Report 92, HMA/SEE, Fairlie Laboratory, Eng., Aug. 1, 1942.

13. *Alternating Current Sensitivity of Recorder Paper,* OSRD WA-317-26, Internal Report 99, HMA/SEE, Fairlie Laboratory, Eng., September 1942.

14. *Split Beam Direction Finding,* R. D. Keynes and R. B. Serle, OSRD WA-647-4, Internal Report 91, HMA/SEE, Fairlie Laboratory, Eng., Apr. 16, 1943.

15. *Eosin Impregnated Recorder Paper,* E. A. Alexander and J. Hawkins, OSRD WA-716-35, Internal Report 133, HMA/SEE, Fairlie Laboratory, Eng., May 20, 1943.

16. *Standard Method of Testing Recorder Paper,* J. Hawkins, OSRD WA-716-36, Internal Report 134, HMA/SEE, Fairlie Laboratory, Eng., May 29, 1943.

17. *Silent Fathometers, General Discussion* (Memorandum), Harry Nyquist, 6.1-NDRC-123. Oct. 20, 1943. Div. 6-634.1-M4

18. *BDI Sample Traces* (External Memorandum), OSRD 3402, NDRC 6.1-sr287-1441, HUSL, Feb. 21, 1944. Div. 6-631.411-M8

19. *Sea Tests of Overhearing of the Secure Echo-Sounding Equipment, SESE Model 21 Aboard the Submarine SS41, Spadefish,* David H. Ransom and Raymond C. Fisher, NDRC 6.1-sr30-1843, Navy Project NS-221, Memorandum SM-251, UCDWR, Aug. 22, 1944. Div. 6-634.2-M2

20. *Single Tone Carrier Telegraphy, Comparison of Masking Effect of Thermal Noise vs. Reverberations,* Kenneth W. Pfleger, 6.1-NDRC-1482, Report 210-KWP-QY, Aug. 28, 1944. Div. 6-560.32-M4

Chapter 9

1. *Laboratory Tests of the Masking of Echoes by Reverberation,* R. S. Alford and W. E. Reid, NDRC 6.1-sr346-2126, BTL, June 8, 1945. Div. 6-560.32-M6

2. *Reverberation Studies at 24 kc,* OSRD 1098, NDRC C4-sr30-401, Navy Project NS-140, Report U-7, Reverberation Group, UCDWR, Nov. 23, 1942. Div. 6-520-M2

3. *Status Report on Task No. 5. Effect of Short Pulse Lengths and Receiver Bandwidth on Echo Ranging,* BTL, July 15, 1944. Div. 6-632.03-M5

4. *Computed Maximum Echo and Detection Ranges for Submarine Echo-Ranging Gear*, William B. Snow and Edward Gerjuoy, NDRC 6.1-sr1131, 1128-1688, Navy Project NS-140, Sonar Analysis Section, CUDWR-NLL, July 1944. Div. 6-570-M2

5. *Echo Recognition Group Training, as developed for use with the echo recognition Monitor Recorder Model 2*, Problem 5H, Report U-325, UCDWR, July 10, 1945. Div. 6-321.4-M1

Chapter 10

1. *The Effect of Doppler on Echo Detection*, OSRD WA-262-10, Internal Report 81, HMA/SEE, Fairlie Laboratory, Eng., May 1942.

2. *Frequency Sensitivity Curves for Type W-621 Headphones*, OSRD W-282-15, Internal Report 89, HMA/SEE, Fairlie Laboratory, Eng., May 21, 1942. Div. 6-560.2-M3

3. *Comparison of Loudspeaker and Telephones for Detectability of Weak Echoes*, H. N. McNair and R. Hall, OSRD WA-2246-5A, Internal Report 178, HMA/SEE, Fairlie Laboratory, Eng., May 8, 1944. Div. 6-560.3-M4

4. *Comparison of Loudspeaker and Telephone for Recognition of Doppler*, OSRD WA-3065-8, Internal Report 199, HMA/SEE, Fairlie Laboratory, Eng., Oct. 17, 1944. Div. 6-560.32-M5

5. *Interim Report on Electronic Own Doppler Nullifier (Mark II)*, A. W. Nolle and W. A. Felsing, NDRC 6.1-sr287-719, HUSL, Mar. 24, 1943. Div. 6-631.31-M5

6. *Automatic Gain Control in Echo-Ranging Systems* (Memorandum for file), F. V. Hunt, NDRC 6.1-sr287-764, HUSL, Apr. 13, 1943. Div. 6-631.13-M1

7. *Description of Devices to be Demonstrated on USS Semmes and USS Galaxy* (Memorandum), NDRC 6.1-sr287-776, HUSL, May 22, 1943. Div. 6-631-M1

8. *Automatic Gain Control in Echo Ranging Systems*, J. Anderson, OSRD WA-854-15-a, Report D-4330, HMA/SEE, Fairlie Laboratory, Eng., July 14, 1943. Div. 6-560.32-M2

9. *Time-Varied Gain for Sonar Equipment* (Completion Report), OSRD 5340, NDRC 6.1-sr287-2076, HUSL, June 15, 1945. Div. 6-631.11-M3

10. *Reverberation Controlled Gain for Sonar Equipment* (Completion Report), F. V. Hunt, OSRD 5415, NDRC 6.1-sr287-2079, HUSL, July 15, 1945. Div. 6-631.12-M5

11. *Sonar Doppler Applications*, F. V. Hunt, OSRD 6558, NDRC 6.1-sr287-2069, HUSL, Nov. 15, 1945. Div. 6-631.3-M5

12. *Frequency-Sensitivity Curves for Brown A Type Headphones*, T. Emmerson, OSRD WA-394-21, Internal Report 114, HMA/SEE, Fairlie Laboratory, Eng., December 1942.

13. *The Doppler Doubler and Square-Law Amplification* (Memorandum for File), W. A. Meyers, Navy Project NS-142, Report M48, UCDWR, Apr. 1, 1943.

14. *Notes on Doppler Enhancer and Indicators*, OSRD WA-389-31e, May 15, 1943.

15. *Tone Duration as a Factor in Pitch Discrimination*, E. G. Wever, NDRC 6.1-sr30-1465, Navy Project NS-97, Memorandum for File M-179, UCDWR, Feb. 16, 1944. Div. 6-560.1-M4

16. *The D-Series of Doppler Drills and Tests, A Report on the Psychological Standards of Auditory Discrimination*, Adelbert Ford, OSRD 3691, NDRC 6.1-sr30-1505, Navy Project NS-97, Report U-206, UCDWR, March 1944. Div. 6-560.32-M3

17. *A Survey of 4003 Audiograms in Relation to the Performance of Sonar Operations at the West Coast Sound School*, Adelbert Ford, Navy Project NS-97, UCDWR, April 1944. Div. 6-311-M10

18. *Echo Doppler Indicator*, J. Warren Horton, OSRD 4511, NDRC 6.1-sr1128-1924, Navy Project NS-142, Completion Report P36/1262, CUDWR-NLL, Dec. 15, 1944. Div. 6-631.34-M5

19. *Comparison of Echo Recognition at 800 and 500 Cycles*, Adelbert Ford and L. J. Cronbach, Report M-312, UCDWR, Apr. 30, 1945.

20. *Doppler Judgment at Low Beat-Frequency Oscillator Settings*, Adelbert Ford, L. J. Cronbach, and D. F. Lovell, Report M-347, UCDWR, Aug. 10, 1945.

21. *Studies of the Recognition of Submarine Echoes*, Report M-431, Sonar Data Division, UCDWR.

CONTRACT NUMBERS, CONTRACTORS, AND SUBJECT OF CONTRACT

Contract Number	*Name and Address of Contractor*	*Subject*
OEMsr-20	The Trustees of Columbia University in the City of New York New York, N. Y.	Studies and experimental investigations in connection with and for the development of equipment and methods pertaining to submarine warfare.
OEMsr-1128	The Trustees of Columbia University in the City of New York New York, N. Y.	Conduct studies and experimental investigations in connection with and for the development of equipment and methods involved in submarine and subsurface warfare.
OEMsr-1131	The Trustees of Columbia University in the City of New York New York, N. Y.	Conduct studies and investigations in connection with the evaluation of the applicability of data, methods, devices, and systems pertaining to submarine and subsurface warfare.
OEMsr-30	The Regents of the University of California Berkeley, Calif.	Maintain and operate certain laboratories and conduct studies and experimental investigations in connection with submarine and subsurface warfare.
OEMsr-346	Western Electric Company, Inc. New York, N. Y.	Studies and experimental investigations in connection with submarine and subsurface warfare.
OEMsr-692	Western Electric Company, Inc. New York, N. Y.	Conduct studies and experimental investigations in connection with the development of listening and detecting systems suitable for surface craft and for submarines.

INDEX

The subject indexes of all STR volumes are combined in a master index printed in a separate volume. For access to the index volume consult the Army or Navy Agency listed on the reverse of the half-title page.

[redacted]

[redacted]

[redacted]

Cancelled

TSDIN FORM 69 A (13 MAR 47)

RESTRICTED

6-8-26

P-33-2-4

13-336-14

ATI- 16504

U.S. Office of Scientific Research and Development

DIVISION: Sciences, General (33) 5/7/62

SECTION: Physics (2)

CROSS REFERENCES: Submarine noise - Aural recognition (90913.38)

ORIG. AGENCY NUMBER: Vol - 9

REVISION

R

U

AUTHOR(S)

now AD-200787

AMER. TITLE: Recognition of underwater sounds

FORG'N. TITLE:

ORIGINATING AGENCY: O.S.R.D., N.D.R.C., Div. 6, Washington, D.C.

TRANSLATION:

COUNTRY	LANGUAGE	FORG'N.CLASS	U. S.CLASS.	DATE	PAGES	ILLUS.	FEATURES
U.S.	Eng.		Restr.	1946	290	231	photos, tables, diagrs, graphs

ABSTRACT

Material presented is a summary of present knowledge of aural recognition of underwater sounds. While the basic theory of the performance of the ear has been kept in mind, a deliberate attempt has been made to focus attention also on the many practical aspects of the information discussed. In part I the relevant facts on the structure and performance of the ear are presented. Part II analyzes the data on the masking of sounds produced by a distant target, while part III deals with the corresponding results for echoes.

lass'fication cancelled
r ch: nged to CONFIDENTIAL
Order Sec. Army by TAG per I 1818.18 dd 27 Nov 53
(ref. ltr same date)
UTH:
y Dorothy S Fischer, OSA
Signature and Grade

Date 1° December 1953

U Authority: Order Sec Army By TAG
PER I 204313 dtd 23 April 58

RESTRICTED

C U

TITLE: The Ear and Hearing, Part I

ATI- 16505

REVISION (None)

AUTHOR(S) : (Not known)

ORIG. AGENCY : Office of Scientific Research and Development, NDRC, Washington, D.C.

ORIG. AGENCY NO. (None)

PUBLISHED BY : (Same)

PUBLISHING AGENCY NO. (None)

DATE	DOC. CLASS.	COUNTRY	LANGUAGE	PAGES	ILLUSTRATIONS
1946	Restr.	U.S.	English	36	table, diagrs, graphs, drwgs

U

ABSTRACT:

Div. 6 Vol. 9

over

A brief account of present knowledge about the structure and behavior of the ear is given. The chief points around which the discussion of the behavior of the ear is developed are: the faintest wanted sounds audible in the presence and absence of interfering sounds; sensations of pitch and loudness; the auditory effects of sounds containing many component frequencies, and the sensed effects produced by varying the furation of the sounds and their components.

~~EO 10501 dd 5 NOV 1953~~

DISTRIBUTION: ~~Copies of this report obtainable from CADO.~~

DIVISION: Aviation Medicine (19)

SECTION: Flight Physiology (3)

16
2,110

SUBJECT HEADINGS: Ear - Physiology (31431.3); Ear - Morphology (31431.2)

ATI SHEET NO.: R-19-3-22

Central Air Documents Office
Wright-Patterson Air Force Base, Dayton, Ohio

AIR TECHNICAL INDEX

RESTRICTED

Conf per auth: Order Sec Army by TAG per
J 1818.18 dd 27 Nov 53 (ref: ltr same date)

U Authority: Order Sec Army By TAG
PER J 204313 dtd 23 Apr 58

TSDIN FORM 69 A (13 MAR 47)

RESTRICTED

R-33-2-4

6-8-25

13-336-14

ATI- 16506

U.S. Office of Scientific Research and Development

AUTHOR(S)

DIVISION: Sciences, General ~~(33)~~ 62

SECTION: Physics (2) 3

CROSS REFERENCES: Submarine noise - Detection (90913.4)

ORIG. AGENCY NUMBER

Vol-9

REVISION

AMER. TITLE: Listening. Part II

FORG'N. TITLE:

ORIGINATING AGENCY: O.S.R.D., N.D.R.C., Div. 6, Washington, D. C.

TRANSLATION:

U

COUNTRY	LANGUAGE	FORG'N.CLASS	U. S.CLASS.	DATE	PAGES	ILLUS.	FEATURES
U.S.	Eng.		Restr.	1946	120	120	tables, diagrs, graphs

over

ABSTRACT

In underwater sound detection, or recognition, the object which either generates or reflects the desired sound is often called the target, and the signal which is directed toward the gear is known as either the target sound or the echo. The major factors determining the possibility of signal detection are the characteristics of the received signal and background, the properties of the gear, and the limitations of the detector. These factors are discussed in sequence.

~~EO 10501 dd 5 NOV 1953~~

T-2, HQ., AIR MATERIEL COMMAND AIR TECHNICAL INDEX WRIGHT FIELD, OHIO, USAAF

RESTRICTED

WF-O-21 MAR 47 22,500

Conf: Auth: Order Sec Army by TAG per
V 1818.18 dd 27 Nov 53 (ref. ltr same date)

U Authority: Order Sec Army By TAG
PER J204313 dtd 23 Apr 58

TSOIN FORM 69 A (13 MAR 47)

RESTRICTED

U.S. Office of Scientific Research and Development

AUTHOR(S)

DIVISION: Sciences, General (33) 62

SECTION: Physics (2) 5

CROSS REFERENCES: Echo ranging (31434.5)

R-33-2-5

6-8-2-5

1,3-3-5-1,4

ATI- 16507

ORIG. AGENCY NUMBER

Vol 9

REVISION

AMER. TITLE: Echo ranging. Part III

FORG'N. TITLE:

ORIGINATING AGENCY: O.S.R.D., N.D.R.C., Div. 6, Washington, D. C.

TRANSLATION:

U U

over

COUNTRY	LANGUAGE	FORG'N.CLASS	U. S.CLASS.	DATE	PAGES	ILLUS.	FEATURES
U.S.	Eng.		Restr.	1946	121	81	photos,tables,diagrs,graphs

ABSTRACT

As in listening to target sounds, a target echo can be heard only if it is sufficiently intense compared to the background of unwanted sounds. In acoustic and reverberation detection, spectrum and amplitude are the significant characteristics of echoes. The background will generally include both noise and reverberation produced by scattering of the pulse. The echo-background mixture should be presented at a loudness level in the neighborhood of 70 phones for best aural results.

~~EO 10501 dd 5 NOV 1953~~

T-2, HQ., AIR MATERIEL COMMAND | AIR TECHNICAL INDEX | WRIGHT FIELD, OHIO, USAAF

RESTRICTED

WF-O-21 MAR 47 22,500

Conf: auth: Sec army by TAG per J1818, 18 dd
27 Nov 53 (ref ltr same date)

U Authority: Order Sec Army By
TAG PER J004313 dtd 23 Apr 58